

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表2002-521860

(P2002-521860A)

(43) 公表日 平成14年7月16日 (2002.7.16)

(51) Int.Cl. ⁷	識別記号	F I	テ-マ-ト (参考)
H 0 4 L 12/56	1 0 0	H 0 4 L 12/56	F 5 K 0 0 2
H 0 4 J 14/08		H 0 4 J 3/00	1 0 0 Z 5 K 0 2 8
H 0 4 B 10/20		H 0 4 B 9/00	Q 5 K 0 3 0
H 0 4 J 3/00			D
			N
		審査請求 有	予備審査請求 有 (全 87 頁)

(21) 出願番号 特願2000-560684(P2000-560684)
(86) (22) 出願日 平成11年7月1日(1999.7.1)
(85) 翻訳文提出日 平成13年1月16日(2001.1.16)
(86) 国際出願番号 PCT/US 99/14979
(87) 国際公開番号 WO 00/04667
(87) 国際公開日 平成12年1月27日(2000.1.27)
(31) 優先権主張番号 09/118, 437
(32) 優先日 平成10年7月17日(1998.7.17)
(33) 優先権主張国 米国 (US)
(81) 指定国 EP(AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), AU, CA, CN, ID, IN, JP, KR, MX, SG

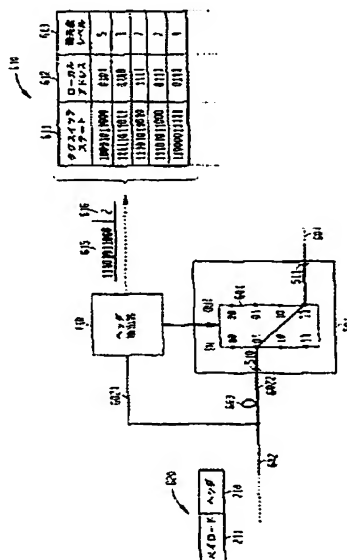
(71) 出願人 テルコーディア テクノロジーズ インコーポレイテッド
アメリカ合衆国 07960-6438 ニュージャージー州 モーリスタウン サウス ストリート 445
(72) 発明者 ジークン チャン
アメリカ合衆国 07733 ニュージャージー州 ホウムデル イースト ローン ドライブ 7
(74) 代理人 弁理士 谷 義一 (外2名)

最終頁に続く

(54) 【発明の名称】 光タグスイッチングを使用した高スループットで低レイテンシの次世代インターネットネットワーク

(57) 【要約】

光ネットワークに応用可能である光シグナリングヘッダ (210) 手法が開示されている。この手法によれば、パケット (620) のルーティング情報は、データペイロード (211) と同じチャネルまたは波長に組み込まれているので、ヘッダ (210) とデータペイロード (211) の両方が、同じ経路および関連遅延を有するネットワークエレメントを通して伝達する。ヘッダ (210) の情報は、データペイロード (211) と十分に異なる特性を持っているので、シグナリングヘッダは、データペイロードに影響されことなく検出可能であり、シグナリングヘッダは、データペイロードに影響を与えることなく除去可能にもなっている。シグナリングルーティング手法は、2つのタイプのプラグアンドプレイモジュールを使用して、モジュール方式によって従来ネットワークエレメントの上に置くことが可能になっている。最初のタイプは、データペイロードがネットワークに出入りする入口点と出口点でヘッダ符号化と復号化を行う。第2のタイプは、ネットワークエレメントの各々でヘッダ検出を行う。



【特許請求の範囲】

【請求項1】 波長分割多重伝送（WDM）ネットワーク内で入力ネットワークエレメントから出力ネットワークエレメントへデータペイロードを伝達するための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有する方法において、

ローカル・ルーティング・ルックアップテーブルを生成し、前記ネットワークエレメントの各々にストアするステップであって、各々のローカル・ルーティングテーブルは、関連する前記ネットワークエレメントの1つを通るローカルルートを決定するステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に光ヘッダを前記データペイロードに追加するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードとヘッダがWDMネットワーク内を伝達するときに前記ネットワークエレメントで前記ヘッダを光学的に判断するステップと、

前記対応するローカル・ルーティングテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記ネットワークエレメントを通る前記データペイロードと前記ヘッダのルート選択をするステップと

を備えたことを特徴とする方法。

【請求項2】 前記光ヘッダは、前記ネットワークエレメントを通る前記光ヘッダと前記データペイロードのルート選択のためのタグスイッチステートを含み、光ヘッダを追加するステップは、前記入力ネットワークエレメントから前記出力ネットワークエレメントへ前記ネットワークエレメントを通る前記光ヘッダと前記データペイロードのルート選択のための該当するタグスイッチステートを判断し、前記光ヘッダに挿入するステップを含むことを特徴とする請求項1に記

載の方法。

【請求項3】 前記光ヘッダは、前記光ヘッダと前記データペイロードが前記ネットワークエレメントを通して伝達するときにルート競合を解決する際に使用する優先度データをさらに含み、ローカル・ルーティングテーブルを生成し、ストアするステップは、優先度レベルを各々のタグスイッチステートに関連付けるステップを含み、前記光ヘッダを追加するステップは、前記データペイロードの該当する優先度データを前記光ヘッダに挿入するステップを含み、選択するステップは、前記優先度データと前記優先度レベルに基づいて前記ローカルルートを選択するステップを含むことを特徴とする請求項2に記載の方法。

【請求項4】 前記光ヘッダを前記データペイロードに追加するステップは、時間的に前記データペイロードより進んだ位置に前記光ヘッダを置くステップを含むことを特徴とする請求項1に記載の方法。

【請求項5】 前記光ヘッダと前記データペイロードは、最初にベースバンドで生成され、前記光ヘッダを前記データペイロードに追加するステップは、前記ベースバンド光ヘッダを、前記ベースバンドデータペイロードの周波数帯域より上の周波数帯域に周波数シフトするステップと、

前記周波数シフトされたベースバンド光ヘッダと前記ベースバンドデータペイロードを結合して複合周波数信号を形成するステップと、

所定の波長の光源を使用して前記複合周波数信号を光変調して、WDMネットワークを通して前記ヘッダと前記データペイロードを伝達する光信号を生成するステップと

を含むことを特徴とする請求項1に記載の方法。

【請求項6】 各々の前記ネットワークエレメントで前記ヘッダを光学的に判断するステップは、

前記光ヘッダをホト検出して検出信号を生成するステップと、

ローカルロッキングオシレータで前記検出信号をロッキングして、ロック信号を生成するステップと、

前記検出信号と前記ロック信号を混合して、ベースバンドの前記ヘッダを表すベースバンド信号を出力するステップと

を含むことを特徴とする請求項5に記載の方法。

【請求項7】 ルート選択をするステップは、前記選択されたルートの競合を解決するステップを含むことを特徴とする請求項1に記載の方法。

【請求項8】 競合を解決するステップは、前記選択されたルートに関連して決定された代替ルートをルート選択するステップを含むことを特徴とする請求項7に記載の方法。

【請求項9】 競合を解決するステップは、前記選択されたルートに使用された波長に関連して決定された代替波長をルート選択するステップを含むことを特徴とする請求項7に記載の方法。

【請求項10】 複数のネットワークエレメントで構成された波長分割多重伝送(WDM)ネットワーク内で入力ネットワークエレメントから出力ネットワークエレメントヘデータペイロードのシーケンスを伝達するための方法であって、前記データペイロードの各々は、所定のフォーマットとプロトコルを有する方法において、

ローカル・ルーティング・ルックアップテーブルを生成し、前記ネットワークエレメントの各々にストアするステップであって、各々のローカル・ルーティングテーブルは、関連する前記ネットワークエレメントの1つを通るローカルルートを決定するステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に光ヘッダを前記データペイロードの各々に追加するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードの各々とそれぞれに対応するヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを示しており、前記データペイロードの各々の前記フォーマットとプロトコルは、それぞれに対応するヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードの各々とそれぞれに対応するヘッダがWDMネットワーク内を伝達するときに前記ネットワークエレメントで前記ヘッダを光学的に判断するステップと、

前記対応するローカル・ルーティングテーブル内の前記ヘッダを調べることに

よって決定された通りに、前記データペイロードの第1とそれに対応するヘッダが前記ネットワークエレメントを通る前記ローカルルートを選択するステップと

、

前記選択されたルートに対応するように前記ネットワークエレメントを通る前記データペイロードの第1とそれに対応するヘッダのルート選択をするステップと、

前記データペイロードの第1で選択された前記ローカルルートを通して前記シケンス中の前記データペイロードの後続する各々のルート選択をするステップと

を備えたことを特徴とする方法。

【請求項11】 ルート選択する各ステップは、前記選択されたルートの競合を解決するステップを含むことを特徴とする請求項10に記載の方法。

【請求項12】 競合を解決するステップは、前記選択されたルートに関連して決定された代替ルートをルート選択するステップを含むことを特徴とする請求項11に記載の方法。

【請求項13】 競合を解決するステップは、前記選択されたルートに使用された波長に関連して決定された代替波長をルート選択するステップを含むことを特徴とする請求項11に記載の方法。

【請求項14】 入力ネットワークエレメントに到着したデータペイロードを、複数のネットワークエレメントで構成された波長分割多重伝送(WDM)ネットワーク上を伝達させるための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有する方法において、

前記データペイロードに関連付けられた光ヘッダを生成するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に前記光ヘッダを前記データペイロードに追加するステップと

を備えたことを特徴とする方法。

【請求項15】 前記光ヘッダと前記ペイロードは、最初にベースバンドで生成され、前記光ヘッダを前記データペイロードに追加するステップは、

前記ベースバンド光ヘッダを、前記ベースバンドデータペイロードの周波数帯域より上の周波数帯域に周波数シフトするステップと、

前記周波数シフトされたベースバンド光ヘッダと前記ベースバンドデータペイロードを結合して複合周波数シグナルを形成するステップと、

所定の波長の光源を使用して前記複合周波数シグナルを光変調して、WDMネットワークを通して前記ヘッダと前記データペイロードを伝達する光信号を生成するステップと

を含むことを特徴とする請求項14に記載の方法。

【請求項16】 複数のネットワークエレメントで構成された波長分割多重伝送（WDM）ネットワーク内で各々の特定のネットワークエレメントの入力から出力にヘッダとデータペイロードを転送するための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有し、ヘッダのフォーマットとプロトコルから独立している方法において、

ローカル・ルーティング・ルックアップテーブルを生成し、前記特定のネットワークエレメントにストアするステップであって、前記ローカル・ルーティングテーブルは、前記特定のネットワークエレメントを通るローカルルートを決めるステップと、

前記ペイロードとヘッダが前記特定のネットワークエレメントの前記入力に到達したときに前記ヘッダを光学的に判断するステップと、

前記ローカル・ルーティングテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが前記特定のネットワークエレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記特定のネットワークエレメントを通る前記データペイロードと前記ヘッダのルート選択をするステップと

を備えたことを特徴とする方法。

【請求項17】 各々の前記ネットワークエレメントで前記ヘッダを光学的

に判断するステップは、

前記光ヘッダをホト検出して検出信号を出力するステップと、

ローカルロッキングオシレータで前記検出信号をロッキングして、ロック信号を生成するステップと、

前記検出信号と前記ロック信号を混合して、ベースバンドの前記ヘッダを表すベースバンド信号を出力するステップと

を含むことを特徴とする請求項16に記載の方法。

【請求項18】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される（a）電気層と（b）光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するためのシステムであって、前記データペイロードは、所定のフォーマットとプロトコルを有するシステムにおいて、

前記ソースデバイスと前記WDMネットワークを結合し、前記データペイロードを前記WDMネットワークに入力する前に光ヘッダを前記データペイロードの前に追加するための光ヘッダモジュールの第1のタイプであって、前記ヘッダは、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している光ヘッダモジュールの第1のタイプと、

各々の前記ネットワークエレメントに付加された光ヘッダモジュールの第2のタイプであって、ローカル・ルーティング・ルックアップテーブルを対応する前記ネットワークエレメントの1つにストアする手段であって、各々のローカル・ルーティングテーブルは、対応する前記ネットワークエレメントの1つを通るルーティング経路を決定する手段と、前記データペイロードとヘッダがWDMネットワーク上を伝達するときに対応する前記ネットワークエレメントの1つで前記ヘッダを光学的に判断する手段と、前記対応するローカル・ルーティングテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが対応する前記ネットワークエレメントの1つを通る前記ローカル

ルートを選択する手段と、前記選択されたルートに対応するように対応する前記ネットワークエレメントの1つを通る前記データペイロードと前記ヘッダのルート選択をする手段とを含む光ヘッダモジュールの第2のタイプとを備えたことを特徴とするシステム。

【請求項19】 前記光ヘッダモジュールの第1のタイプの他方は、WDMネットワークをデスティネーションデバイスに結合しており、前記光ヘッダモジュールの第1のタイプは、前記デスティネーションデバイスに配達する前に前記データペイロードから前記ヘッダを除去する手段を含むことを特徴とする請求項18に記載のシステム。

【請求項20】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される（a）電気層と（b）光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するための光ヘッダモジュールであって、前記データペイロードは、所定のフォーマットとプロトコルを有する光ヘッダモジュールにおいて、前記光ヘッダモジュールは、前記ソースデバイスと前記WDMネットワークを結合し、前記光ヘッダモジュールは、前記データペイロードに関連付けられた光ヘッダを生成する手段であって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している手段と、前記データペイロードを前記入力ネットワークエレメントに入力する前に前記光ヘッダを前記データペイロードに追加する手段とを含むことを特徴とする光ヘッダモジュール。

【請求項21】 前記光ヘッダは、前記ネットワークエレメントを通る前記光ヘッダと前記データペイロードのルート選択のためのタグスイッチステートを含み、光ヘッダを追加する手段は、前記ネットワークエレメントを通る前記光ヘッダと前記データペイロードのルート選択のための該当するタグスイッチステートを判断し、前記光ヘッダに挿入する手段を含むことを特徴とする請求項20に記載のシステム。

【請求項22】 前記光ヘッダを前記データペイロードに追加する手段は、時間的に前記データペイロードより進んだ位置に前記光ヘッダを置く手段を含むことを特徴とする請求項20に記載のシステム。

【請求項23】 前記光ヘッダと前記データペイロードは、最初にベースバンドで生成され、前記光ヘッダを前記データペイロードに追加する手段は、

前記ベースバンド光ヘッダを、前記ベースバンドデータペイロードの周波数帯域より上の周波数帯域に周波数シフトする手段と、

前記周波数シフトされたベースバンド光ヘッダと前記ベースバンドデータペイロードを結合して複合周波数信号を形成する手段と、

所定の波長の光源を使用して前記複合周波数信号を光変調して、WDMネットワークを通して前記ヘッダと前記データペイロードを伝達する光信号を生成する手段と

を含むことを特徴とする請求項20に記載のシステム。

【請求項24】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される（a）電気層と（b）光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイス向けに送られるデータペイロードを伝達するための光ヘッダプロセッサであって、前記データペイロードは、所定のフォーマットとプロトコルを有する光ヘッダプロセッサにおいて、前記光ヘッダプロセッサは、各々の前記ネットワークエレメントと関連付けられ、前記光ヘッダプロセッサは、

ローカル・ルーティング・ループアップテーブルを生成し、対応する前記ネットワークエレメントの1つにストアする手段であって、各々のローカル・ルーティングテーブルは、対応する前記ネットワークエレメントの1つを通るルーティング経路を決定する手段と、

前記データペイロードとヘッダがWDMネットワーク内を伝達するときに対応する前記ネットワークエレメントの1つで前記ヘッダを光学的に判断する手段と、

前記対応するローカル・ルーティングテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが対応する前記ネットワークエレメントの1つを通る前記ローカルルートを選択する手段と、

前記選択されたルートに対応するように対応する前記ネットワークエレメントの1つを通る前記データペイロードと前記ヘッダのルート選択をする手段とを備えたことを特徴とする光ヘッダプロセッサ。

【請求項25】 各々の前記ネットワークエレメントで前記ヘッダを光学的に判断する手段は、

前記光ヘッダをホト検出して検出信号を出力する手段と、

前記検出信号をロッキングしてロック信号を生成するローカルロッキングオシレータと、

前記検出信号と前記ロック信号を混合して、ベースバンドの前記ヘッダを表すベースバンド信号を出力する手段と

を含むことを特徴とする請求項24に記載のヘッダプロセッサ。

【請求項26】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される（a）電気層と（b）光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するためのシステムであって、前記データペイロードは、所定のフォーマットとプロトコルを有し、前記ネットワークは、前記ネットワークエレメントに結合され、前記ネットワークを通る回線交換ルートを決定するネットワークマネージャを含み、各々の前記ネットワークエレメントは、（i）スイッチングデバイスと、（ii）前記ネットワークマネージャに応答して、前記ネットワークマネージャからの入力に基づいて前記スイッチングデバイスを制御して、前記WDMネットワークを通る回線交換ルーティング経路を確立する回線交換コントローラとを含むシステムにおいて、

前記ソースデバイスと前記WDMネットワークを結合し、前記データペイロードを前記WDMネットワークに入力する前に光ヘッダを前記データペイロードの前に追加するための光ヘッダモジュールの第1のタイプであって、前記ヘッダは、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している光ヘッダモジュールの第1のタイプと、

前記ネットワークマネージャと前記回線交換コントローラにตอบสนองし、前記スイッチングデバイスに結合されている光ヘッダモジュールの第2のタイプであって、前記ネットワークマネージャから与えられたローカル・ルーティングテーブルを各々のネットワークエレメントにストアする手段であって、各々のローカル・ルーティングテーブルは、各々のネットワークエレメントを通るルーティング経路を決定する手段と、前記データペイロードと前記ヘッダが前記WDMネットワーク内を伝達されるときに各々の前記ネットワークエレメントで前記ヘッダを光学的に判断する手段と、前記対応するローカル・ルーティングテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを選択する手段と、前記回線交換コントローラからの入力と前記ローカル・ルーティングテーブルを処理して前記スイッチングデバイスを制御することによって、前記選択されたルートに対応するように各々の前記ネットワークを通る前記データペイロードと前記ヘッダのルート選択をする手段とを備えたことを特徴とするシステム。

【請求項27】 前記スイッチングデバイスの前に置かれていて、前記スイッチングデバイスへの前記データペイロードと前記ヘッダの配達を、あらかじめ決めたインターバルで遅延させる手段をさらに含むことを特徴とする請求項26に記載のシステム。

【請求項28】 各々の前記ネットワークエレメントで前記ヘッダを光学的に判断する手段は、前記WDMネットワーク上を伝達する前記ヘッダをベースバンドのヘッダに復調するデマルチプレクサを含み、

選択する手段は、前記デマルチプレクサにตอบสนองして、前記ベースバンドのヘッダに含まれるルート情報を判断するための高速メモリを含み、

ルートを選択する手段は、前記ローカル・ルーティングテーブルに結合され、前記高速メモリと前記回線交換コントローラにตอบสนองして、前記スイッチングデバイスを制御するためのタグスイッチコントローラを含むことを特徴とする請求項26に記載のシステム。

【請求項29】 複数のネットワークエレメントを含む波長分割多重伝送（

WDM) ネットワークで構成される (a) 電気層と (b) 光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するためのシステムであって、前記データペイロードは、所定のフォーマットとプロトコルを有するシステムにおいて、

前記ソースデバイスと前記WDMネットワークを結合し、前記データペイロードを前記WDMネットワークに入力する前に光ヘッダを前記データペイロードの前に追加するための光ヘッダモジュールの第1のタイプであって、前記ヘッダは、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している光ヘッダモジュールの第1のタイプと、

各々の前記ネットワークエレメントに付加された光ヘッダモジュールの第2のタイプであって、ローカル・ルーティング・ルックアップテーブルを対応する前記ネットワークエレメントの1つにストアする手段であって、各々のローカル・ルーティングテーブルは、対応する各々の前記ネットワークエレメントの1つを通るルーティング経路を決定する手段と、前記データペイロードとヘッダが前記WDMネットワーク上を伝達するときに対応する前記ネットワークエレメントの1つで前記ヘッダを光学的に判断する手段と、前記対応するローカル・ルーティングテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが対応する前記ネットワークエレメントの1つを通るローカルルートを選択する手段と、前記選択されたルートに対応するように対応する前記ネットワークエレメントの1つを通る前記データペイロードと前記ヘッダのルート選択をする手段と、同じローカルルートを有する後続の連続するヘッダの各々に対して前記選択されたルートを維持する手段とを含む光ヘッダモジュールの第2のタイプと

を備えたことを特徴とするシステム。

【発明の詳細な説明】

【0001】

(発明の背景)

(1.発明の分野)

本発明は、光通信システムに関し、さらに具体的には、高スループットで低レイテンシのネットワーク・トラフィックを受け入れる光システムに関する。

【0002】

(2.従来技術の説明)

光波長分割多重伝送 (WDM) 技術の最近の研究進歩により、伝送帯域幅が、既存の商用ネットワークよりも数倍も高くなっているネットワークの開発が促進されている。このようなスループット向上は、それだけを取り上げると印象的であるが、国民的発意 (national initiatives) を含む、新しいアプリケーションをサポートするという要求条件を満足できる次世代の超高速ネットワークを提供する次世代インターネット (NGI) を実現するためには、ネットワーク・レイテンシの低減化も、同じように達成されなければならない。この目的に向かって、現在の研究努力は、WDM光パケットスイッチング技術を利用した超低レイテンシのインターネットプロトコル (IP) の開発に集中しており、高スループットで低レイテンシという2つの目標を達成することを約束している。この研究努力は、将来性がある反面、この2つの目標を実現するまでには至っていない。

【0003】

このようなIP/WDMネットワークを実現するには、チャレンジする要求条件がいくつかある。第1に、NGIネットワークは、既存インターネットと相互運用され、プロトコル間の衝突 (conflicts) を回避しなければならない。第2に、NGIネットワークは、超低レイテンシを提供するだけでなく、パケット交換 (つまり、バーストの) IPトラフィックと回線交換WDMネットワークの両方を活用しなければならない。第3に、NGIネットワークが、シグナリングとデータペイロードの間で同期をとる必要がなければ好都合である。最後に、望まれている目標は、NGIネットワークが種々のプロトコルとフォーマットのデータトラフィックを受け入れ、複雑な同期化やフォーマット変換を必要としないで、IPシグナルと非IPシ

グナルの両方を送受信できるようにすることである。

【0004】

(他の作業との比較)

R. E. Wagner他の論文「MONET: Multi-Wavelength Optical Networking」で報告され、Journal of Lightwave Technology, Vol. 14, No. 6, June 1996に公表されている、マルチ波長光ネットワークは、光ネットワークにおける重要なマイルストーンをいくつか示しており、全国規模のファイバ距離にわたって12を超える再構成可能なネットワークエレメントを通したマルチ波長の透過伝送が示されている。しかし、このネットワークは回線交換であり、バーストラフィックを受け入れるには非効率である。要求からスイッチングまでの代表的な接続セットアップ時間は数秒であり、これは、ネットワーク制御および管理 (NC&M) とハードウェアの能力に制約があるためである。MONETプログラムにおいて効率を向上するための最近の努力は、「ジャストインタイム・シグナリング」方式に集中している。この方式は、組み込み (embedded) 1510 nm NC&Mシグナリングを利用し、このシグナリングは、推測遅延時間だけデータペイロードに先行している。この推測は、シグナリングヘッダとペイロードのスイッチングを同期させるために、各々のネットワーク構成ごとにすべての波長について正確に行われなければならない。

【0005】

本発明によれば、光パケットヘッダは、パケットペイロードデータと同じ波長に乗せて伝搬される。この方法によると、ヘッダとペイロードを同期させるという問題がなくなる。さらに、各中間光スイッチでの光遅延を適切に用いると、光遅延がスイッチで直接的に吸収されるので、初期バースト遅延を推測する必要がなくなる。これは、経路上の各スイッチでの遅延が前もって分かっている必要があり、総遅延の計算にその遅延を含める必要のあるジャストインタイム・シグナリングと著しく異なっている。最後に、接続時間の要求と実際の接続の達成において浪費される時間が少なくなっている。MONETでは、数秒の遅延があるのに対し、本発明の主題によれば、遅延は最小限になり、その制約となるのは、各スイッチでの実際のハードウェア・スイッチング遅延だけである。現在のスイッチン

グ技術では、わずか数ミリ秒の遅延が実現されているが、この遅延は、将来には、さらに短縮される可能性がある。この短時間の遅延は、スイッチを利用する各ネットワークエレメントで光ファイバ・ディレイラインを使用することによって吸収することができる。本発明の主題によれば、可能な限りのレイテンシ低減化が、ハードウェアの基本的制約のぎりぎりまで達成されているが、他の方法ではレイテンシをこれ以上低減化することは不可能である。

【0006】

光ネットワークテクノロジー・コンソーシアム（ONTC）の結果は、Chang他の論文「Multiwavelength Reconfigurable WDM/ATM/SONET Network Testbed」で報告され、Journal of Lightwave Technology, Vol. 14, No. 6, June 1996に公表されている。ONTCプログラムのフェーズI（155Mb/s、4波長）とフェーズII（2.5Gb/s、8波長）は共に、マルチホップATMベースのネットワーク上に構成されている。このようなATMベースのアーキテクチャによると、大きなオーバーヘッドが付加され、シングルホップネットワークの可能性はないが、パケット／ヘッダシグナリングは、等時性（isochronous）のATMセル自体を利用することによって可能にしている。このNC&M情報の通信は、同じ光波長を通して行われるので、本発明の手法による場合と同じ利点が得られる潜在性をもっている。しかし、本発明の手法によれば、ATMベースのシグナリングに比べて、いくつかの顕著な利点が得られる。第1に、本発明の手法によれば、電気シグナルに変換し、パケットをバッファに置いておかなくても、シングルホップコネクションが得られる。第2に、本発明の手法によれば、余分のオーバーヘッドが除去されるので、帯域幅の利用効率がさらに向上にする。第3に、本発明の手法によれば、厳格に透過的で、超低レイテンシなコネクションが得られる。

【0007】

ARPA（Advanced Research Projects Agency）後援の全光ネットワーク（AON）コンソーシアムの結果は、I. P. Kaminow他の論文「A Wideband All-Optical WDM Network」で報告され、IEEE Journal on Selected Areas of Communication, Vol. 14, No. 5, June 1996に公表されている。AONプログラムは、実際には2つの部分からなっている。前掲の論文に報告されているWDMと、同じ号の姉妹論文

に報告されているTDMである。最初に、AONプログラムのWDM部分について先に論じ、そのあとでTDM部分について論じることとする。

【0008】

AONアーキテクチャは、3レベルのサブネットワーク階層構造になっており、これはコンピュータネットワークに見られるLAN、MAN、およびWANのそれに似ている。AONは、光端末（OT）間の3つの基本サービスを提供している。すなわち、Aサービス、Bサービス、およびCサービスである。Aは、透過回線交換サービス、Bは、時間スケジュール透過TDM/WDMサービス、Cは、シグナリングで用される非透過データグラムサービスである。Bサービスは、250 msecフレームが使用され、フレームごとに128スロットをもつ構造を使用している。あるスロットまたはスロットグループ内では、変調レートとフォーマットの選択は、ユーザの自由に任されている。AONアーキテクチャに実装されているBサービスは、本発明の主題となっている、WDM上のIP（IP over WDM）に最も近くなっている。しかし、CサービスのNC&MシグナリングとBサービスのペイロードは分離されているので、シグナリングヘッダとペイロード間の同期をとるには慎重さが要求される。この要求条件は、250ミリ秒フレームが使用され、フレームごとに128スロットで、ビットレートが任意であると、さらに厳しくなる。同期化は、ビットレベルで行う必要があるだけでなく、この同期化はネットワーク全体にわたって達成される必要がある。スケーラビリティ（scalability）と相互運用性（interoperability）は、これらが、要求されるネットワーク同期化と歩調していないので極めて困難である。本発明の主題によれば、同期化は要求されず、既存のIPおよび非IPトラフィックと相互運用され、スケーラビリティが得られる。

【0009】

TDMでは100Gb/sビットレートを努力目標としている。原理的には、この種の超高速TDMネットワークは、真にフレキシブルなオンデマンド帯域幅（bandwidth on demand）を、100Gb/sのバーストレートで提供する潜在性をもっている。しかし、この種の高ビットレートシステムの背後には、主にファイバの非線形性、拡散、および偏光低下に関連する重大な技術上のチャレンジが潜んでいる。ソリトン技術によると、これらの困難性のいくつかは緩和できるが、それでも、ネット

パケットの非常に正確な同期化、つまり、数ピコ秒までの同期化が要求される。さらに、パケットとペイロードは同一ビットレートになっている必要があるため、ビットレート透過サービスを提供することが困難になっている。本発明による主題では、同期化が要求されず、100Gb/s技術に頼らないので、透過サービスが得られるようになっている。

【0010】

Cisco Corporationは、タグスイッチング (Tag-Switching) をベースとするプロダクトを最近発表した。なお、Cisco社のタグスイッチングの概要説明は、WWW (world-wide-web) サイト (<http://www.cisco.com/wrap/public/732/tag/>) にアクセスすることで入手できる。Cisco社の (電子的) タグスイッチングでは、ルータとスイッチからなるネットワークを通過するパケットに、ラベル、すなわち「タグ」を割り当てている。従来のルータネットワークでは、各パケットは、その最終デスティネーション (宛先) に向かうパケットの、次のホップを判断するために各ルータに処理させなければならない。(電子的) タグスイッチング・ネットワークでは、タグは、デスティネーション・ネットワークまたはホストに割り当てられている。そして、パケットは、ネットワークを通るようにスイッチングされ、各ノードは、各パケットを処理するのではなく、タグをスワップするだけである。(電子的) タグスイッチング・ネットワークは、(電子的) タグスイッチ (従来のルータまたはスイッチのどちらか) のコアから構成され、タグスイッチは、ネットワークの周辺上の (電子的) タグ・エッジルータに接続している。(電子的) タグ・エッジルータとタグスイッチは、標準ルーティングプロトコルを使用して、ネットワークを通るルートを判断している。そのあと、これらのシステムは、ルーティングプロトコルによって生成されたテーブルを使用して、タグ情報を割り当て、タグ配信プロトコルでそのタグ情報を配信している。タグスイッチとタグ・エッジルータは、タグ配信プロトコル情報を受信し、転送データベースを構築する。このデータベースは、特定のデスティネーションを、そのデスティネーションに関連付けられたタグおよびタグがそこから到達可能であるポートと対応付けている。

【0011】

タグ・エッジルータは、タグネットワーク上を転送するパケットを受信すると、ネットワーク層ヘッダを分析し、適用可能なネットワーク層サービスを実行する。そのあと、パケットのルートをそのルーティングテーブルから選択し、タグを付け、パケットを次ホップのタグスイッチに転送する。

【0012】

タグスイッチは、タグ付きパケットを受信し、ネットワーク層ヘッダを再分析することなく、タグだけに基づいてパケットをスイッチングする。パケットは、ネットワークの出口点 (egress point) にあるタグ・エッジルータに到達し、そこでタグが取り除かれ、パケットが配達される。Cisco社が(電子的)タグスイッチングの発表を行った後、IETF (Internet Engineering Task Force) は、ベンド中立の(電子的)タグスイッチング機能を標準化し、ルータおよびATMスイッチを含むスイッチに実装するためのMPLS (Multi-protocol Label) を勧告した。

【0013】

Cisco社の(電子的)タグスイッチングの特徴のいくつかは、本発明の主題になっている光タグスイッチング (Optical Tag Switching) に類似し、その特徴は、パケットルーティングに要求される処理を単純化するという類似の目標を目的としている。主な相違点は次のとおりである。第1に、光タグスイッチングは、タグとデータペイロードが共に光の形になっているという意味で、純粹に光学的である。各プラグアンドプレイモジュール (本発明システムのコンポーネントの1つ) は、光タグを検知するが、実際のパケットはネットワークから出るまでは、光から電気への変換が行われない。Cisco社の(電子的)タグスイッチングはすべてが電氣的であり、各ルータで電子的検出、処理、および再送信を各パケットに適用している。第2に、本発明の光タグスイッチングは、可能な限りの低レイテンシを達成し、バッファを利用することに頼っていない。電子的タグスイッチングは、処理が電子的に行われ、バッファリングが電子的に行われるため、レイテンシがはるかに大きくなっている。第3に、本発明の光タグスイッチングは、経路偏向 (path deflection) および/または波長変換を利用してパケットの競合 (contention) に起因するブロッキングを解決しているのに対し、電子的タ

グスイッチングは、レイテンシが増加することを犠牲にして競合の解決を達成する手段として、電子バッファリングだけを利用し、パフォーマンスは、パケットサイズに大きく左右されている。本発明が対象とするパケットの長さは任意である。最後に、本発明の光タグスイッチングは、厳格に透過のネットワークを達成しており、そこでは、任意のフォーマットとプロトコルのデータは、正しい光タグが付いている限り、ルーティングが可能になっている。従って、データのフォーマットは、任意のビットレートのデジタルにすることも、アナログにすることも、FSK (frequency-shifted-keying) にすることも可能である。電子的タグスイッチングでは、データペイロードは、電子タグと同じ所定のデジタルビットレートになっている必要があるが、これは、ルータがこれらを電子的にバッファリングしなければならないからである。

【0014】

本発明の背景となるもう1つの代表的な技術は、セッション偏向仮想回線プロトコル (Session Deflection Virtual Circuit Protocol SDVC) と呼ばれているもので、これは偏向ルーティング手法 (deflection routing method) をベースにしている。Proceedings on IEEE Globecom 85, pp 255-261, December 1985に掲載された、N. F. Maxemchukの「The Manhattan Street Network」というタイトルの論文は、2つのパケットが同じデスティネーションに送られようとしているとき、一方のパケットは、優先出力リンク (preferred output link) 用にランダムに選択できるが、他方のパケットは非優先リンク (non-preferred link) に「偏向」されることを論じている。このことは、パケットは最短経路でない経路を偶然にとることを意味している。本発明が利用している偏向手法では、最優先経路に送られるパケットを「ランダム」に選択するのではなく、パケットの優先度を調べ、優先度の高いパケットを優先経路にルーティングされるように送信するようにしている。パケットは、その優先度が低ければ偏向される。しかし、「経路偏向 (path deflection)」と「波長偏向 (wave length deflection)」の両方が利用されている。経路偏向は、光パケットが同じ波長の次の優先経路にルーティングされるだけである点で、従来のSDVCに類似している。波長偏向によると、光パケットを最優先経路にルーティングできるが、波長が異なっている。

この波長偏向は、ネットワークエレメントで波長変換することにより達成される。部分的に制限波長変換 (limited wave length conversion) が利用されているのは、発信元周波数が与えられているとき、すべての波長がデスティネーションの波長として利用できるとは限らないからである。波長偏向によると、経路遅延を増加することなく、波長コンテンションに起因するブロッキングを解決することができる。経路偏向と波長偏向を併用すると、パケット競合を解決するために十分に大きな付加的な接続性 (connectivity) が得られるが、部分的な波長変換の度合は、ブロッキングレートが上昇し始めると増加することになる。ネットワークのこのようなスケーラビリティとフレキシビリティは、従来のSDVCでは取り上げられていない。

【0015】

(発明の概要)

本発明は、光ネットワークに応用可能である、ユニークな光シグナリングヘッダ手法を利用している。パケットルーティング情報は、データペイロードと同じチャネルまたは波長に組み込まれているので、ヘッダとデータ情報は共に、同じ経路および関連した遅延で、ネットワークを伝達されるようになっている。しかし、ヘッダルーティング情報は、データペイロードとは異なる特性をもっている。シグナリングヘッダは、データペイロードに影響されることなく検出することができると共に、シグナリングヘッダは、データペイロードに影響することなく取り除くことができるようになっている。本発明の主題によれば、このユニークなシグナルルーティング手法は、2種類の「プラグアンドプレイ (Plug-and-play)」モジュールを追加することによって、従来のネットワークエレメントの上にモジュール方式で置くことが可能になっている。本発明の主題によれば、「発明の背景」の個所で説明している他の手法の欠点と制約が解消されると共に、光ネットワークがもつ全能力を活用できるという利点がある。

【0016】

本発明の一側面である広い意味での方法によれば、複数のネットワークエレメントで構成された波長分割多重伝送システムにおいて、データペイロードを入力ネットワークエレメントから出力ネットワークエレメントに伝達するための方法

は、そのデータペイロードが所定のフォーマットとプロトコルをもっていることを前提として、(a) ローカル・ルーティング・ルックアップテーブルを生成し、ネットワークエレメントの各々にストアするステップであって、各々のローカル・ルーティングテーブルは、関連するネットワークエレメントの1つを通るローカルルートを決定するステップと、(b) データペイロードを入力ネットワークエレメントに入力する前に光ヘッダをデータペイロードに追加するステップであって、ヘッダは、フォーマットとプロトコルを有し、データペイロードとヘッダが各々のネットワークエレメントを通るローカルルートを示しており、データペイロードのフォーマットとプロトコルは、ヘッダのフォーマットとプロトコルから独立しているステップと、(c) データペイロードとヘッダがWDMネットワーク内を伝達するときにネットワークエレメントでヘッダを光学的に判断するステップと、(d) 対応するローカル・ルーティングテーブル内のヘッダを調べることによって決定された通りに、データペイロードとヘッダがネットワークエレメントを通るローカルルートを選択するステップと、(e) 選択されたルートに対応するようにネットワークエレメントを通るデータペイロードとヘッダのルート選択をするステップとを備えている。

【0017】

本発明の別側面である広い意味でのシステムによれば、システムは、複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される (a) 電気層と (b) 光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するためのシステムであり、データペイロードは、所定のフォーマットとプロトコルを有する。システムは、(i) ソースデバイスとWDMネットワークを結合し、データペイロードをWDMネットワークに入力する前に光ヘッダをデータペイロードの前に追加するための光ヘッダモジュールの第1のタイプであって、ヘッダは、データペイロードとヘッダがネットワークエレメントを通るローカルルートを示しており、データペイロードのフォーマットとプロトコルは、ヘッダのフォーマットとプロトコルから独立している光ヘッダモジュールの第1のタイプと、(ii) 各々のネットワークエレメントに付加された光ヘッダモジュール

ルの第2のタイプであって、ローカル・ルーティング・ルックアップテーブルを対応するネットワークエレメントの1つにストアする手段であって、各々のローカル・ルーティングテーブルは、対応するネットワークエレメントの1つを通るルーティング経路を決定する手段と、データペイロードとヘッダがWDMネットワーク上を伝達するときに対応するネットワークエレメントの1つでヘッダを光学的に判断する手段と、対応するローカル・ルーティングテーブル内のヘッダを調べることによって決定された通りに、データペイロードとヘッダが対応するネットワークエレメントの1つを通るローカルルートを選択する手段と、選択されたルートに対応するように対応するネットワークエレメントの1つを通るデータペイロードとヘッダのルート選択をする手段とを含む光ヘッダモジュールの第2のタイプとを含んでいる。

【0018】

本発明によれば、多数の特徴と利点が得られる。そのいくつかを示すと、(1) レイテンシが極めて低く、その制約となるのはハードウェアの遅延だけであること、(2) 高スループットとオンデマンド帯域幅が、マルチ波長ネットワークと光タグスイッチングを結合することによって得られること、(3) 優先度に基づくルーティングによって、優先度の高いデータグラムまたはパケットが高スループットを得られるようにしたこと、(4) 従来のWDMから本発明の光タグ交換WDMへのネットワークのアップグレードがスケーラブルで、モジュール方式で行われること、(5) 長いデータグラム、連続するパケット、および非連続のパケットでさえも、ルーティングが効率的に行われること、(6) マルチプレクサやファイバなどの光コンポーネントをコスト効率的に利用すること、(7) マルチベンダ環境内で相互運用性があること、(8) ネットワークエレメントのアップグレードがグレースフルに、ステップバイステップで行われること、(9) どのフォーマットの、どのプロトコルのデータも、透過的にサポートされること、(10) コミュニケーションが高サービス品質であること、などである。

【0019】

(詳細な説明)

本発明による基本的原理の理解を容易にし、以下の説明の中で使用されている

用語を紹介するために、まず概要を説明し、そのあとで例示の実施形態について説明することにする。

【0020】

(概要)

本発明は、NGIアプリケーションのデータが大きなブロックからなるとき、低レイテンシで高スループットの、コスト効率のよいオンデマンド帯域幅を実現するネットワークに関するものである。ネットワークをコスト効率的にかつ相互運用可能にアップグレードすることは、既存のWDMネットワークエレメントの上にポータブル「プラグアンドプレイ」モジュールを置き、「WDM光タグスイッチング」と呼ばれているものを実施することによって実現されている。なお、「光タグスイッチング」も同じ意味で使用されている。本発明は、あらゆる観点で、NGIネットワークのハードウェアとソフトウェアの両方に影響を与えている。この中には、アーキテクチャ、プロトコル、ネットワーク管理、ネットワークエレメントの設計、およびイネープリングテクノロジー (enabling technologies) が含まれている。

【0021】

ネットワークによって実施される技法とそのネットワークを実現するための付属回路は、以上から理解されるように、インバンド光シグナリングヘッダによるバーストの間にルーティング経路を動的に生成するものと定義されている、WDM光タグスイッチングと呼ばれる手法によって生まれたものである。データパケットは、各パケットのインバンドWDMシグナリングヘッダを使用して、WDMネットワークを通るようにルート選択される。スイッチングノードでは、シグナリングヘッダが処理され、ヘッダとデータペイロードは、(1) 既存のフロー状態のコネクションを通して即時に転送されることも、(2) バーストの間に経路をセットアップして、ヘッダとデータペイロードを処理するようにすることも可能である。WDMタグスイッチングによると、ルーティングとスループットが高効率化され、パケットが光レベルでルーティングされることを保つために要求されるIPレベルのホップ数を1ホップに減少することが可能になり、その管理は、ルーティング情報を作成し、維持しているNC&Mによって行われる。

【0022】

図1は、中間層130で光層120と電気層110を結合することによって得られる、汎用ネットワーク100の光層120と電気層110の相互関係を示す図である。電気層110は、簡略化のために、2つの従来のIPルータ111と112から構成されているものとして示されている。光層120は、ネットワークエレメントまたはノード121～125から構成されているものとして示されている。中間層130は、IPルータ112をネットワークエレメント122に結合する従来のATM/SONETシステム131を示している。層130の一部としてヘッダネットワーク132も示されているが、これは、本発明によれば、IPルータ111をネットワークエレメント121に結合している。図1には、完全な相互運用性と再構成可能性をもつ、全国規模の透過WDMをベースとするバックボーンネットワーク上のネットワーク100の配置が図示されている。ここで強調しておきたいことは、図1のエレメントは本発明による実施形態の一例であり、従って、エレメント111は、別の実施形態では、ATMルータにすることも、スイッチにすることも可能である。

【0023】

次に、図2を参照して説明すると、図1の光層120が詳細に示されており、この中には、本発明によれば、ネットワークエレメント121～125から構成された光ネットワーク201で、高速のコネクションをセットアップするための基本的手法が含まれている。セットアップでは、そのあとに続くデータペイロード211に対して光シグナリングヘッダ210が使用される。この手法は、回線交換をベースとするWDMの利点と、パケット交換をベースとするIPテクノロジーの利点の両方を兼ね備えている。新規のシグナリングヘッダ210は、光シグナルヘッダ210の形で追加されており、マルチ波長トランスポート環境では、各波長内をインバンドで搬送される。光シグナリングヘッダ210は、パケットのソース、デスティネーション、優先度、およびパケットの長さなどの、ルーティング情報と制御情報とを収めているタグであり、データペイロード211に先行して光ネットワーク内を伝達される。各WDMネットワークエレメント121～125は、光シグナリングヘッダ210を検知し、コネクションテーブル（後述する

）を調べ、クロスコネクション（cross-connection）、アッド（add）、ドロップ（drop）、ドロップ・コンティニュー（drop-and-continue）などの、必要なステップを実行する。コネクションテーブルは、NC&M 220とWDMネットワークエレメント121～125が継続的に通信し合うことによって絶えず更新されている。光シグナリングヘッダ210のあとに続くデータペイロード211は、コネクションによって確立された各ネットワークエレメントの経路を通るようにルーティングされる（後述する）。図2の構成によれば、図2にTで示すように、光シグナリングヘッダ210とデータペイロード211の間の時間遅延を管理する必要がないのは、各ネットワークエレメント内でコネクションセットアップのために要求される短時間に対して必要な光遅延が、その間に置かれたファイバの遅延を通して各ネットワークエレメントで与えられるからである。さらに、データペイロードのフォーマットとプロトコルは、ヘッダのそれから独立している。すなわち、あるネットワークが与えられているとき、ヘッダのフォーマットとプロトコルは事前に決まっているのに対し、データペイロードのフォーマットとプロトコルはヘッダのそれと同じにすることも、異なったものにすることも可能になっている。

【0024】

各デスティネーションは、「コスト」を最小限にする優先経路と関連付けられている。図2に示すように、ソース123からデスティネーション122までの経路全体には、カスケード状の経路201と202が含まれており、両経路とも波長WPを利用している。このコストは、総伝達距離、ホップの数、およびトラフィック負荷に基づいて計算される。優先波長は、オリジナル波長がデフォルトになっている。例えば、経路202上の優先波長はWPになっている。デフォルト波長にある、この優先経路がすでに別のパケットによって占有されていれば、ネットワークエレメント121は、同じ優先経路を通る、利用可能な代替波長WAがあるかどうかを即時に判断する。この代替波長は、ネットワークエレメント121での制限波長変換によって得られる波長から選択できる波長の1つになっていなければならない。選択できる波長がないために、最優先経路を通るようにパケットを転送できないときは、次の優先経路が選択される（経路偏向）。例えば、図

2に示すように、カスケード状の経路203と204は代替経路を表すことができる。この時点で、優先波長は、デフォルトのオリジナル波長WPに戻される。このデフォルト波長が再びすでに占有されていれば、同じプロセスで代替波長を探すことができる。図2では、経路203は同一波長WPをもつ代替経路であり、経路204は代替波長WAを使用している代替経路である。経路を組み合わせたものが存在しないため、波長偏向によってパケットのトランスポートを可能にすることは起こり得ないケースであるが、そのようなことが起こった場合には、ネットワークエレメント121は、優先度の低いパケットをドロップすることを決定することになる。言い換えれば、発信元波長の優先通路を通る新規パケットの転送は、すでに優先経路を占有している、優先度の低い他方のパケットをドロップすることによって行われる。

【0025】

ネットワークエレメント121～125は、「プラグアンドプレイ」と呼ばれる2つのタイプのモジュールで強化され、従来の回線交換WDMネットワークエレメント121～125に、パケットスイッチング能力を持たせることによって、バーストラフィックの処理を効率化している。そこでは、シグナリングヘッダは符号化されてIPパケット上に載せられ、必要なときに削除される。

【0026】

第1タイプの「プラグアンドプレイ」モジュールは、図1に電気／光モジュール132として示されているが、図3にはブロック図で示されている。概念的には、モジュール132は、スタンドアロンエレメントであるが、実際には、モジュール132は図3に示すように、ネットワークエレメント121と統合化されている。モジュール132は、ネットワークエレメント121のコンプライアント・クライアント・インタフェース (compliant client interface CCI) 310とIPルータ111の中間に置かれ、ヘッダエンコーダ321を使用して光シグナリングヘッダ210を符号化して、ネットワーク内に追加されたパケットの上に載せ、またネットワークからドロップされたパケットから光シグナリングヘッダ210を、ヘッダリムーバ322を使用して除去する。

【0027】

一般的に、符号化／除去モジュール132は、IPトラフィックがWDMネットワークに入ったり、出たりするインタフェースとなっている個所に置かれ、これはネットワークエレメントのクライアント・インタフェースとIPルータの間に位置している。クライアント・インタフェースは、CCIタイプにすることも、非コンプライアント・クライアント・インタフェース (non-compliant client interface NCI) タイプにすることも可能である。これらのインタフェースの個所で、ヘッダエンコーダ321は、IPシグナルがネットワーク201に送り込まれるとき、デスティネーションと他の情報を収めている光ヘッダ210を、データペイロード211の前に置く。光ヘッダ210は、光変調器（後述する）によって光ドメイン内で符号化される。シグナリングヘッダリムーバ322は、クライアント・インタフェースからドロップされた光シグナルのヘッダ210を検出し、電氣的なIPパケットをIPルータ111に引き渡す。

【0028】

さらに具体的に説明すると、モジュール132は、IPルータ111から電気シグナルを受け取り、その電気シグナルを必要とするコンプライアント波長光シグナルに変換し、光ヘッダ210をパケット全体の前に置く。モジュール132は、NC&M220と通信し、NC&M220が要求していれば、データをバッファに置いてからデータを光学的に変換する。モジュール132は、クライアント・インタフェースの波長に整合された波長を持つ光送信器（後述する）を採用している。（後述するが、ここで参考のために触れておくと、モジュール132が図4のNCI404とも互換性があるのは、波長適合 (wavelength adaptation) がNCIで行われるからである。しかし、NCI波長適合と光ヘッダをもつIPシグナルのビットレート互換性は、事前に設定されていなければならない。

【0029】

図4は、第2タイプの「プラグアンドプレイ」モジュールである光エレメント410を示しているが、これは、各々のWDMネットワークエレメント121～125と関連付けられている。ここでは、説明の便宜上、エレメント121が示されている。モジュール410は、従来のネットワークエレメント回線スイッチコントローラ420と、従来のスイッチングデバイス430との間に置かれている

。モジュール410は、タップファイバ経路 (tapped fiber path) 404～406からモジュール410へ供給された、いずれかのファイバ401～403上を伝達された各シグナリングヘッダ210からの情報を検出する。モジュール410の働きは、テーブルルックアップを高速化し、スイッチングデバイス430へのシグナリングを高速化することである。スイッチコントローラ420は、ネットワークエレメントを制御するために使用されている従来の「クラフトインタフェース (craft interface)」と同じ働きをする。しかし、この場合には、このスイッチコントローラ420の目的は、NC&M220から回線交換シグナリングを受け取り、タグスイッチコントローラ410にどのコマンドを送るべきかを、優先度に基づいて判断することである。従って、タグスイッチコントローラ410は、ネットワークエレメント回線スイッチコントローラ420から回線交換制御シグナルのほかに、各シグナリングヘッダ210から引き出された情報も受信し、回線交換方式にするか、タグスイッチ制御方式にするかを統合的に選択する。スイッチングデバイス430を構成するスイッチも、スイッチングを高速化する。スイッチングデバイス430への入力経路401～403に置かれているファイバ415、416、または417に要求される遅延は、遅延がシグナリングヘッダを読み取り、テーブルルックアップを完了し、スイッチングを行うために要する総時間より大になるようになっている。概算であるが、2kmファイバでは、処理時間は10ミリ秒になる。エレメント121～125で表され、スイッチングデバイス430を含んでいるWDMネットワークエレメントのタイプとしては、波長アッドドロップ・マルチプレクサ (Wave length Add-Drop Multiplexer WADM)、波長選択クロスコネクト (Wave length Selective Crossconnect WSXC)、および制限波長変換能力もつ波長インターチェンジ・クロスコネクト (Wave length Interchanging Crossconnect WIXC) がある。

【0030】

オペレーション時には、モジュール410は、経路401～403上に現れた光シグナルの小部分をタップすることによって、各シグナリングヘッダ210に入っている情報を検出し、モジュール410にストアされたコネクションテーブルを調べた後、スイッチングデバイス430に合った該当コマンドを判断する。

ファイバ遅延は経路401～403に置かれるので、ヘッダ210とデータペイロード211をもつパケットが、実際のスイッチングが行われた後だけスイッチングデバイス430に到着するようにする。このファイバ遅延は、ヘッダ検出、テーブルルックアップ、およびスイッチングに関連する遅延に固有のものであるので、ファイバ415～417のファイバ約2kmの遅延は、約10マイクロ秒で行うことができるのが代表的である。

【0031】

ネットワークエレメント121～125では、データペイロード211の光から電気への変換も、電気から光への変換も行われないので、コネクションは完全に透過になっている。IPルーティングでは、ビットレートが多数であり、プロトコルが低レベルであるため、必要とされる異種インタフェースの数が増加し、その結果、ルータのコストが上昇するのに対し、WDMタグスイッチングによるルーティングは、ビットレートから見えなくなっている。一例として、ネットワークエレメント121～125による光ルーティングでは、1.28Tb/secのスループットを達成することで（ファイバ当たり32波長で、波長当たり2.5Gb/secである16×16クロスコネクトスイッチングデバイス430の場合）、これは現行のギガビットルータのどれよりも大幅に向上している。

【0032】

各ネットワークエレメント121～125をNC&M220と結合すると、適応性のあるルーティングプロトコルが実施される。このルーティングプロトコルは、次のような機能を実行する。(a) 通信回線のステート、予測トラフィック、遅延、容量利用など、ルーティングストラテジに関係するネットワークパラメータを測定し、(b) ルーティング計算のために測定情報をNC&M220に転送し、(c) NC&M220でルーティングテーブルを計算し、(d) ルーティングテーブルを各ネットワークエレメント121～125に配布し、各ネットワークエレメントでパケットルーティングの判断をさせる。NC&M220は、各ネットワークエレメントからネットワークパラメータ情報を受信し、ルーティングテーブルを定期的に更新し、そのあと(e) エレメント111などのIPルータからのコネクション要求をNC&M220に転送し、(f) NC&M220からのルーティング情報を各ネットワークエ

レメント121～125に転送し、光シグナリングヘッダ210に入力させる。

【0033】

パケットは、各パケットのシグナリングヘッダ210に入っている情報を使用して、ネットワーク201を通るルートが選択される。パケットがネットワークエレメントに到着すると、シグナリングヘッダ210が読み取られ、パケットは、(a) タグ・ルーティング・ルックアップテーブルに従って、選択された新しい該当出方路ポートにルーティングされるか、あるいは(b) ネットワークエレメント内の既存のタグスイッチング発コネクションを経由して即時に転送される。後者は「フロー・スイッチング (flow switching)」と呼ばれ、光タグスイッチングの一部としてサポートされている。フロー・スイッチングは、バーストモードトラフィックが大量であるとき使用される。

【0034】

タグスイッチ・ルーティング・ルックアップテーブルは、フロー・スイッチングステートがセットアップされていないとき、ネットワークエレメントを通る光パケットのルート選択を高速化するために、ネットワークエレメント121～125に組み込まれている。光シグナリングヘッダ210で伝達されるコネクションセットアップ要求は、各ネットワークエレメント内のタグスイッチ・ルーティング・ルックアップテーブルと突き合わせて高速に比較される。場合によっては、最も効率よくシグナルをルーティングする最適コネクションは、すでに占有されていることがある。この起こり得るコネクションルックアップテーブルは、代替波長割り当てまたはシグナルをルーティングする代替経路をすでに提供している構成にもなっている。代替波長の数を制限すると（少なくとも1つに）、ブロッキングが起こる可能性が大幅に減少する。代替波長ルーティングによると、伝達遅延とホップの数も最適な場合と同じになるので、複数のパケットを順序付けするときの問題がなくなる。代替経路ルーティングによると、遅延とホップ数が増加する可能性があるので、パケットが非常に多数のホップを通るようにルーティングされる可能性は、パケットの信号対雑音比を光学的にモニタすることによって除去されている。第2の経路または波長が利用可能でない場合には、出方路リンクでの競合は、先着順サービス (first-come, first service) 方式または

優先度方式で解決することができる。情報は正規のIPルータに提示された後、上位層プロトコルによって検査され、必要時には再送が使用される。

【0035】

(ルーティングの例)

図5に図示されているのは、米国のいくつかの大都市にあるエンドユーザ間でパケットを伝達するための、WDM回線交換バックボーンネットワーク500の例である。まず、ネットワーク500は、従来ではどのように運用されているかについて説明し、本発明によるWDM光タグスイッチングのオーバーレイは、そのあとで説明することにする。

【0036】

図5を参照して説明すると、そこでは、ニューヨーク市はネットワークエレメント501がサービスし、シカゴはネットワークエレメント502がサービスし、ロス・アンジェルスはネットワークエレメント504がサービスし、ミネアポリスはネットワークエレメント507がサービスするものとしており、その他の都市も同様である。(ネットワークエレメントは、以下の説明ではノードと呼ぶこともある)。さらに、NC&M220は、物理層の光監視チャネルを経由する、すべてのネットワークエレメントへの論理コネクション(ネットワークエレメント501へのチャネル221およびネットワークエレメント507へのチャネル222のように、破線で示されている)をもっている。NC&M220とネットワークエレメント501～507との間は、継続的通信状態にある。NC&M220は、(a) 各ネットワークエレメントの一般状態(例えば、動作状態にあるか、緊急時のシャットダウン状態にあるか)、(b) 各ネットワークエレメントから提供される光波長(例えば、ネットワークエレメント501は、波長W1をもつ光ファイバ媒体531と、波長W2をもつ光ファイバ媒体532とによってサービスされ、これらの光ファイバ媒体は、それぞれネットワークエレメント502(シカゴ)とネットワークエレメント505(ボストン)に接続しているものとして示されている)、(c) 波長によってサービスされるポート(例えば、エレメント501のポート510は、着信クライアント・インタフェース伝達パケット520に関連付けられ、ポート511はW1に関連付けられ、ポート512はW2に関連付けられ、

他方、エレメント502のポート513は、W1に関連付けられている)に関する情報を定期的に要求し、受信する。

【0037】

以上のように、NC&M220は、ネットワークエレメントによる着信パケットトラフィックを搬送するルートを定式化するために必要なグローバル情報を、瞬時にストアしている。従って、NC&M220は、例えば、グローバル・ルーティングテーブル形体のルーティング情報を定期的に判断し、監視チャネル221、222などを使用してグローバル・ルーティングテーブルをエレメントの各々にダウンロードする。グローバル・ルーティングテーブルは、ネットワークエレメントのポートが、ある種の通信リンクを構築する構成になっている。例えば、NC&M220は、ニューヨーク市からロス・アンジェルスへの光ファイバリンク（それぞれネットワークエレメント501と504）が現在要求されていることを、トラフィック要求と統計に基づいて判断し、リンクは、エレメント501のポート511をネットワークエレメント502内のポート513に結合しているW1、エレメント502のポート514をエレメント503のポート515に結合しているW1、およびエレメント503のポート516をエレメント504のポート517に結合しているW2で連続的に構成されることになる。そのあと、ネットワークエレメント501（ニューヨーク市）に着信し、デスティネーションがネットワークエレメント504（ロス・アンジェルス）となっている入力パケット502は、この確立されたリンク上を即時にルーティングされる。ネットワークエレメント504では、伝達パケットはクライアント・インタフェースポート518から出力パケット521として配達される。

【0038】

同じように、エレメント506と507（それぞれセント・ルイスとミネアポリス）間の専用経路は、ネットワークエレメント506と502間ではW3を使用し、エレメント502と507間ではW2を使用して確立されているものとして示されている。

【0039】

このように、グローバル・ルーティングテーブルに基づいて生成されたリンク

はその非融通性に特徴がある。つまり、NC&M 220がリンクを確立するコネクションを判断し、リンクの接続性情報をダウンロードし、各ネットワークエレメントの入力ポートと出力ポートを確立するのに数秒を要している。各リンクは、回線交換コネクションの特性をもっている。つまり、このリンクは、基本的には、長いインターバルに対して固定接続（permanent connection）か専用経路に、あるいは「パイプ」になっており、NC&M 220だけが通常オペレーションでリンクを切り離し、再確立することが可能になっている。この種の専用経路の利点は、発信元（origin）とデスティネーションをもち、すでに確立された専用経路にマッピングされているトラフィックが、どのセットアップも必要としないで即時にルーティングできることである。他方、専用経路は、専用経路が使用されるのが全体の時間のわずかなだけ（例えば、セットアップ期間全体にわたって20%～50%）であるという点で非効率になる可能性があり、ほとんどの場合がそうである。さらに、各ネットワークエレメントに組み込まれ、入力ポートと出力ポートを相互接続しているスイッチングデバイス430（図4参照）は、入力ポートと出力ポートの数が有限数に制限されている。セント・ルイスからミネアポリスまでのリンクが必要になり、ニューヨークからロス・アンジェルスまでのリンクに、すでに割り当てられているポート（例えば、ネットワークエレメント502のポート514）が使用されるように上記シナリオが変更されていれば、NC&M 220が応答し、その応答に見合うようにグローバル・ルーティングテーブルを変更するまでに時間遅延が生じる。

【0040】

次に、本発明の原理による主題を、上述してきた説明に上乗せされるように拡張された例について説明することにする。最初に、「タグスイッチステート（tag-switched state）」と名付けたパラメータが導入され、このパラメータがルーティングでどのように使用されるかについて説明する。そのあとで、タグスイッチステートがどのように生成されるかについて説明する。タグスイッチステートが生成されると、光タグスイッチングが行われる。

【0041】

NC&M 220は、さらに、クライアント・インタフェースから各ネットワークエ

メントに着信する各パケットに、タグスイッチステートを割り当てるような構成になっている。タグスイッチステートは、プラグアンドプレイモジュール132によって追加されているが、本発明の目的上、タグスイッチステートは、ヘッダ210（図2参照）と同等の働きをする。タグスイッチステートは、NC&M220によって計算され、ローカル・ルーティングテーブルの形で各ネットワークエレメント501～507にダウンロードされる。図6を参照して説明すると、ネットワークエレメント501とその組み込みスイッチ601が図示されている。ヘッダ210とペイロード211で構成されたパケット620を搬送する着信光ファイバ602も、遅延ループ（delay loop）603と共に示されているが、ペイロード211は、このケースでは、図5からのパケット520になっている。ファイバ6022は、遅延されたバージョンのパケット620をネットワークエレメント501に配達する。ファイバ602上に現れた光エネルギーの一部は、ファイバ6021でタップされた後、光モジュール410に入力され、そこで着信パケット620が処理され、ヘッダ210が検出される。パケット620のヘッダ210は、参照符号615で示しているタグスイッチステート 11101011000からなるものとして示されている。図6には、ルックアップテーブル610も示されているが、これは、2つのカラム、つまり、「タグスイッチステート」（カラム611）と「ローカルアドレス」（カラム612）から構成されている。パケット620の特定タグスイッチステートは、ルックアップテーブル610で相互参照され、着信パケットのルーティングが判断されるようになっている。このケースでは、パケット620のタグスイッチステートは、ルックアップテーブル610の4行目のエントリになっている。このタグスイッチステートに対応するローカルスイッチアドレスは“0111”になっているが、その意味は次のとおりである。最初の2桁（バイナリ）は、着信ポートを示し、次の2桁（バイナリ）は出力ポートを示している。このケースでは、4入力、4出力スイッチの例では、着信パケットは、入力ポート“01”から出力ポート“11”へのルートが選択されるので、スイッチ601はそうようにスイッチングされる（図示）。ファイバ遅延603によって遅延が与えられた後、ファイバ6022上の着信パケットは、スイッチ601を経由してファイバ604上を伝達される。

【0042】

タグスイッチステートがどのように使用されるかは、上述した通りである。次に、タグスイッチステートがどのように生成されるかについて説明する。NC&M 220は、この場合も定期的に、対応する各ネットワークエレメントを通るルートを選択／スイッチングするためのローカルルックアップテーブル群（ネットワークエレメント501ではテーブル610）をコンパイルし、そのあと各ルックアップテーブルは、対応するネットワークエレメントにダウンロードされる。各ルックアップテーブルの生成には、NC&M 220がネットワーク500のグローバル知識をもっていることが考慮される。例えば、ネットワーク501への着信パケット620がネットワーク504に向けられ（この場合も、ニューヨークからロス・アンジェルスへ）、ポート501が着信ポート"01"に関連付けられていてファイバ602にサービスしており、出力ポート511が送出ポート"11"に関連付けられていてファイバ604にサービスしている場合には、NC&M 220は、ルックアップテーブル610に該当エントリ（すなわち、4行目）を生成し、テーブル610をネットワーク510にダウンロードすることができる。そこで、パケット520は、電気／光モジュール132によって処理され、ヘッダ210がパケット520に追加されて増補パケット620が作られるとき、NC&M 220は、ダウンロードされたローカル・ルーティングテーブルの知識だけでなく、パケット520に組み込まれていて、モジュール132から得られたデスティネーションアドレスの知識も有しているので、NC&M 220は、該当タグスイッチステートをヘッダ210、この場合には 11101011000 として追加するようにモジュール132に指示することができる。

【0043】

以上から容易に理解されるように、タグスイッチステートパラメータを使用してパケットを処理することは、その性質上バーストになっている。すなわち、スイッチ601が、着信タグスイッチステートを処理するようにセットアップされると、スイッチ601は、フローステートを処理する以前のステートに戻されることになる。例えば、スイッチ601は、パケット620の到着前に入力ポート01を出力ポート10に相互接続していれば、データ処理の後のステート011

0 に戻されることになる（これは、例えば、パケットのトレーラから判断される）。回線交換経路がタグスイッチステート経路と同じであることも、当然に考えられるが、その場合には、スイッチ601を通るローカルルートを変更しなくても、タグスイッチステートが処理される。しかし、スイッチ601を一時的に変更する必要があるときは、基礎となる回線交換トラフィックがあれば、ルートを選択し直すことも、再送することも可能である。

【0044】

以上に説明したように、タグスイッチングによると、ネットワークエレメントがデータパケット全体を調べなくても、パケットをデスティネーション向けにルーティングすることが可能になる。新規のシグナリング情報、つまり、タグは、光シグナルヘッダ210の形で追加され、これはマルチ波長トランスポート環境では各波長内をインバンドで搬送される。このタグスイッチングは、パケット単位で行われるのが通常である。しかし、パケットは非常に多数が同一デスティネーション向けに順次に転送されるのが一般的である。このことが特に当てはまるのがバーストデータであり、そこでは大きなデータブロックは、多数のパケットに分割されて転送されている。このような場合、各々の特定ネットワークエレメントが各タグを慎重に検査し、ルーティング経路を判断することは非効率的である。むしろ、ソースからデスティネーションへの「仮想回線 (virtual circuit)」をセットアップした方が効率的である。各パケットのヘッダ210は、仮想回線を継続するか、終了するかを通知するだけである。これはフローステートコネクション (flow state connection) と呼ばれている。このようなエンドツーエンド・フローステート経路が確立された後、ネットワークエレメント内のプラグアンドプレイモジュールは、切り離しが必要になるまでこのフローステートコネクションを維持している。切り離しが行われるのは、一連のパケットがエンドに到着したか、はるかに優先度の高い別のパケットがこのフローステートコネクションの切り離しを要求したときである。

【0045】

優先度から見た本発明の側面も、図6に示されている。ローカルルックアップテーブルには、「優先度レベル」(カラム613)があり、これはタグスイッチ

ングステートに割り当てられた優先度を示している。また、ヘッダ210には、番号2で示された優先度データが追加されている（参照符号616）。テーブル610の「タグスイッチステート」カラムの4行目と5行目は共に、ローカルアドレスが0111になっている。以前のデータパケットが5行目のエントリを使用して、例えば、仮想回線またはフロー・スイッチングステートを確立していて、今は別になっているパケットが、カラム611の4行目に示すように処理されれば、優先度の高いデータ（"4"ではなく"2"になっている。1が最も優先度が高い）に優先権があるので、仮想回線は終了することになる。

【0046】

（実施例の詳細説明）

WDMタグスイッチング上の超低レイテンシIPを達成するためには、各光スイッチでの光ヘッダの処理は、光パケットが実際に伝送されている間、最小限に保たれていなければならない。この目的を達成するために導入されたのが、光WDMタグスイッチングを実行するための新規のシグナリングアーキテクチャとパケット伝送プロトコルである。

【0047】

シグナリングとパケット伝送プロトコルでは、低速で複雑なIPルーティング機能は、超高速WDMスイッチング機能から切り離されている。この切り離しは、エンドツーエンドのルーティング経路のセットアップを通して達成されるが、その頻度は非常に少なくなっている。ソースからデスティネーションにIPパケットを送信するには、次のようなステップで行われる。

【0048】

(a) エンドツーエンドのルーティングセットアップ。ここでは、IP層のソフトウェアは、ネットワークエレメントとNC&M220間のシグナリングプロトコルを起動し、IPパケットのためのエンドツーエンドのルーティング経路をセットアップする。このステップでは、ルーティング経路上のWDMネットワークエレメントが、以後のパケット転送をサポートするようにも構成される。実際のパケット伝送期間中に、光ヘッダに挿入される光タグスイッチングのタグも判断される。

【0049】

(b) 光パケット伝送。ここでは、光パケットが到着すると、ローカルヘッダ処理がトリガされ、この処理では、特に、パケットを次のホップまで転送するための出力ポートが、光ヘッダ内の光タグに基づいて調べられる。

【0050】

ルーティング経路セットアップでは、一般に低速で複雑な手続きであるルーティング機能が起動されるが、これはパケット伝送処理に先だって行われるので、伝送レイテンシを決定することは、クリティカル・パスにならない。

【0051】

(ルーティング経路セットアップ)

ルーティング経路セットアップ期間に、WDMパケットスイッチの内部接続テーブルは、タグスイッチルックアップテーブルで増大され、関係するパケット転送情報を収容している。具体的には、超低レイテンシとハードウェアの簡略化を達成するために、本発明の方式により得られるタグスイッチステートは、フロー経路上を一定に保たれている。例えば、タグスイッチ割り当て手法には、次のものがある。

【0052】

(1) デスティネーションをベースとするフロータグの割り当て。この方式では、デスティネーション、例えば、適当なデスティネーションIPアドレスプレフィックスは、次のホップを調べるときにタグスイッチステートとして使用することができる。光ヘッダを変更する必要があるだけでなく、偏向ルーティングを行うときにも同じヘッダが使用できる。

【0053】

(2) ルートをベースとするフロータグの割り当て。この方式では、割り当てられたタグスイッチステートは、タグスイッチステートセットアップフェーズで動的に計算されるエンドツーエンドルートを指している。この方式の利点は、個別タグスイッチステートごとに、サービス品質 (Quality of Service) 要求条件を満たすように特殊化できることである。

【0054】

(スイッチングの衝突の解決)

現在、実現性のある光バッファ技術は存在していない。このことは、従来のバッファリング手法では、スイッチングの衝突を処理できないことを意味している。上述したように、本発明の実施形態では、光フィルタに実装されている光遅延を利用して、この時間的遅延の間にスイッチングが行われるようにしているが、従来のIPルータにおける電氣的バッファがそうであるように、競合の解決を達成するようになっていない。スイッチング競合を解決するために、本発明によれば、次の3つの方法が使用されている。

【0055】

(a) 制限波長インターチェンジ (limited wavelength interchange)。この方法では、パケットは同じ経路を通るルートが選択されるが、異なる波長になっている。この波長変換は、コンテンション回避だけを目的に利用されているので、全波長チャネルのいずれかに変換する能力を、ネットワークエレメントにもたせる必要がない。むしろ、全波長チャネルの一部を変換できれば十分である。この波長変換は、シグナリングヘッダとデータペイロードの両方を変換する。パケットが受ける波長変換が多すぎると、信号の正確さが低下するので、そのことを避ける必要がある。1つ可能なポリシーは、変換を1つだけにすることである。これは、光ヘッダ内のオリジナル波長を符号化することで容易に実施することができる。このようにすると、中間WDMスイッチは、オリジナル波長で搬送されるときだけ変換を許すことになる。

【0056】

(b) 制限偏向ルーティング (limited deflection routing)。この方法では、パケットを隣接スイッチングノードに偏向させ、そこからデスティネーションに転送することを可能にしている。この場合も、パケットが繰り返し偏向されると、信号の品質低下の原因になるだけでなく、ネットワーク帯域幅が浪費されることになるので、それを避けるための注意が必要である。これを解決する方法は、「タイムスタンプ」フィールドを光ヘッダに記録し、記録したタイムスタンプが最大の制限値より古くないときにだけ偏向が行われるようにすることである。

【0057】

(c) 優先パケット回避 (prioritized packet preemption)。この方法では、

新着パケットは、その着信パケットの優先度が高いとき、現在伝送中のパケットに優先することができる。その目的はすべてのパケットに公平性を保証し、再送パケットが最終的に配達されることを保証することである。この方式では、各パケットは、この場合も、その光ヘッダに記録されたタイムスタンプ・フィールドを有し、古い方のパケットが新しい方のパケットよりも優先度が高くなっている。さらに、再送パケットは、オリジナルパケットのタイムスタンプをもっている。このようにすると、パケットが「古く」となると、優先度が高くなり、最終的には、必要であればそのデスティネーションに向かう通路を優先使用できるようになる。

【0058】

なお、上記3方式のいずれも、光ヘッダは、ネットワーク内を移動するとき、常に一定のままになっている。これは、光スイッチングハードウェアを高速かつ単純化するという目的に合致している。また、これらの方式を組み合わせることも可能である。

【0059】

(ルーティングプロトコル)

NGIと同規模のネットワークでは、ルーティング判断を集中化することは実現不可能であるので、このアプローチは、分散意思決定まで一般化する必要がある。階層アドレッシングとルーティングは、IPルーティングの場合と同じように使用されている。新しいコネクションが要求されると、NC&M 220は、WDM経路がWDMベースのネットワーク内のそのペア（ソース、デスティネーション）に用意されているかどうかを判断する。そうであれば、パケットは、即時にその（1ホップIPレベルの）経路上に送出される。そのような経路が用意されていなければ、NC&M 220は、最初のWDMネットワークエレメント用の初期出方路リンクと、新トラフィックを搬送する波長を判断する。この判断は、新コネクションの要求時に、ネットワーク内に残っているコネクションに基づいて行われる。その後、NC&M 220は、該当するプロトコルを通してシグナリングを使用し、関係情報を初期WDMネットワークエレメントに転送し、シグナリングヘッダに入れられる。初期出方路リンクが判断された後、ルーティング判断の残りは、光シグナリングヘッ

ダ情報に従って個別のNEで行われる。この方法によると、各スイッチングノードに置かれたルーティングテーブルと、必要とされるシグナリングヘッダ処理が、相対的に小さく保たれることが保証される。また、この方法によると、ネットワークは、スイッチングノードとネットワークユーザの点でスケーリングが容易化される。なお、複数のWDMサブネットを1つに相互接続し、各サブネットに独自のNC&Mをもたせることも可能である。

【0060】

経路が決定されたとき、WDM NE内では、光スイッチは、ノードを通る各パケットの持続時間の間そのステートにセットし、(i) そのあとデフォルトのステートに戻すことも（光タグスイッチングと呼ばれる）、(ii) 有限の短時間の間そのステートにセットしておくことも（フロー・スイッチングと呼ばれる）可能である。前者のケースでは、ルーティングは正規のパケット単位で行われる。システムリソースは、送信される情報があるときだけ専用され、パケットが終了すると、これらのリソースは、別のパケットへの割り当てのために解放される。後者のケースは、バーストモードのトラフィックが大量であるとき使用される。このケースでは、WDM NEは、NEに到着する後続パケットの光シグナリングヘッダからフローステートタグを読み取るだけでよく、スイッチングデバイスを切り替えなくても、パケットが同じデスティネーション向けであることを確かめ、光タグスイッチングによって以前に確立された、NEを通る既存コネクションを通してペイロードを転送することができる。

【0061】

パケットは、各パケットのシグナリングヘッダに入っている情報を使用して、ネットワークを通るルートが自己選択される。パケットがスイッチングノードに到着すると、シグナリングヘッダが読み取られ、パケットが既存のフローステートコネクションを通して即時に転送されるか、あるいはルーティングテーブルに従って新しい該当出方路ポートが選択される。各ノードのルーティングテーブルは、波長ごとに存在している。別のパケットとの競合のために、パケットが選択された出方路ポートをたどって行くことができないときは（選択された出方路ファイバが解放されていない）、このルーティング方式では、同じ出方路ポートに

対して別の波長を割り当てることが試みられる（その結果、シグナルはスイッチングノード内で波長変換を受けることになる）。選択された出方路ポートで利用できる、適格な波長が他になければ、別のテーブルから別の出方路ポートが選択されることになる。このテーブルには、二次的な（優先度で見たとき）出方路リンクがリストされている。

【0062】

本発明による手法では、このルーティングプロトコルは、偏向ルーティング方式に似ている（「発明の背景」の個所参照）。そこでは、セッションは、優先経路にたどって行くことができないとき、他の出方路リンク（優先度で見たとき）に偏向される。パケットは、継続的に偏向されることは許されていない。従来のルーティングプロトコルでは、特定数のホップの後、セッションをブロックするためにホップカウントが使用されている。新規方式では、スイッチングノードでヘッダ再生成が許されていない場合には、ホップカウント手法は使用できない。その代わりに、光シグナリングヘッダの特性（つまり、シグナリングヘッダのSNR）を調べて、パケットをドロップすべきかどうかを決定することが可能になっている。

【0063】

（WDM層のIPルーティングアルゴリズム）

ルーティングテーブルを判断するために、NC&M 220によって使用される手法は、最小コストの経路を通るルートを選択して、ソースからデスティネーションにパケットを転送する最短経路アルゴリズム（shortest path algorithm）に基づいている。長さ、容量利用率、ホップカウント、平均パケット遅延といった、各ルート上の具体的コスト基準は、異なるネットワークで使用することができる。このルーティング機能の目的は、パフォーマンスを良好化すると共に（ネットワーク内を通過する平均遅延を低くすること）、高スループットを維持することである。最小コストスパニングツリー（minimum cost spanning tree）が生成され、その生成のたびに異なるノードがルートとなっている。これらのツリーから得られた情報は、各スイッチングノードでルーティングテーブルをセットアップするために使用することができる。上述した偏向ルーティングが実装されていれ

ば、k-最短経路手法を使用すると、多数の潜在的ルーティング経路を利用することができる。この手法によると、2つ以上の最短経路が探し出され、これらの経路はコスト順にランク付けされている。この情報をスイッチングノードのルーティングテーブルに入力すると、最小コスト経路に対応する出方路リンクが最初に考慮され、コストが大の経路に対応するリンクは、偏向ルーティングを実装するために使用される二次ルーティングテーブルに入力される。

【0064】

(プラグアンドプレイモジュールの説明)

本発明は、WDMネットワークエレメントに付加される2つのタイプのプラグアンドプレイモジュールをベースにしている。これらのプラグアンドプレイモジュールを導入すると、光タグスイッチング能力が既存回線交換ネットワークエレメントに付加されることになる。

【0065】

図3では、ヘッダエンコーダ321とヘッダリムーバ322は、ハイレベルのブロック図形式で示されていた。図7と図8は、それぞれ、エンコーダ321とリムーバ322の両方を示す詳細概略図である。

【0066】

図7に示すように、IPパケットまたはデータグラムは、マイクロプロセッサ710で処理され、タグスイッチングのための各光シグナリングヘッダ210が生成される。光シグナリングヘッダ210とオリジナルIPパケット211とは、ベースバンドでマイクロプロセッサ710から出力される。シグナリングヘッダ210は、ローカルオシレータ730を利用して、RFミキサ720で混合される。ミキサ720からの混合されたヘッダとオリジナルパケット211とは、コンバイナ740で結合され、コンバイナ740の出力は、レーザ750を変調ソースとしている光モジュレータ760で光波長チャネルに符号化される。

【0067】

図8に示すように、ネットワークエレメントからドロップされた光チャネルは、フォト検出器810によって検出され、増幅器820によって電氣的に増幅される。通常、フォト検出器810と増幅器820のどちらも、周波数応答は、データ

ペイロードだけをカバーし、ローカルオシレータ730から得られる光シグナリングヘッダのRF搬送周波数はカバーしていない。ローパスフィルタ830は、残留RF搬送波があれば、それを除去する。フィルタ830の出力は基本的には、発信元ネットワークエレメントから発信元IPルータによって送出され、ネットワーク内をトランスポートされ、別のネットワークエレメントの別のIPルータによって受信されるオリジナルIPパケットである。

【0068】

図9のブロック図900は、タグスイッチング・シグナリングヘッダ210とデータペイロード211の両方を搬送する光シグナル901を、ベースバンド電気シグナリングヘッダ902に変換するために、図4のプラグアンドプレイモジュール410によって実行される検出プロセスの要素を示している。最初に、光シグナル901は、ホット検出器910によって検出され、ホット検出器901の出力は増幅器920によって増幅され、ハイパスフィルタ930を通して、光シグナリングヘッダ210を搬送する高周波成分だけが残される。RFスプリッタ940は、シグナルをフィードバックロック機能をもつローカルオシレータ950に与える。ローカルオシレータ950からの出力とスプリッタ940からの出力は、ミキサ960で混合される。すなわち、高周波搬送波は、フィルタの出力から減算され、タグスイッチング・シグナリングヘッダ210上の情報だけが残される。このプロセスでは、フィードバックロック機能をもつローカルオシレータ950は、正確な周波数、位相、および振幅をもつローカルオシレーションを得るために使用され、高周波成分は、このローカルオシレータシグナルとタグスイッチング・シグナリングヘッダをシグナリング高周波搬送波とミキシングするときにゼロにされる。ローパスフィルタ970は、ミキサ960の出力に結合され、ベースバンド・シグナリングヘッダ210を電気出力信号902として出力する。

【0069】

図10の回路図は、図4の実施形態の詳細を示す例である。図10において、各ヘッダ検出器1010、1020、...、1030、...、または1040は、それぞれ経路1001、1002、1003、および1004上にあつて

、デマルチプレクサ1005、1006、1007、および1008によって処理された光入力を構成している各波長からの情報を処理する。各デマルチプレクサの例は、図9の回路900に示されている。処理された情報は、各波長別にグループ化される。従って、例えば、高速メモリ1021は、ある波長が与えられているとき、ヘッダ検出器1010からのリード1011とヘッダ検出器1030からのリード1014に現れたシグナルを、入力として受け入れる。各高速メモリ1021～1024は、内容アドレスメモリ（content-addressable memory）のように、対応するタグスイッチコントローラ1031～1034への入力となる働きをしている。各タグスイッチコントローラ1031～1034は、図4のネットワークエレメントスイッチコントローラ420から回線交換制御シグナルも受信する。各タグスイッチコントローラは、コントローラ420から与えられる回線交換コントロールと、対応する高速メモリから与えられるタグスイッチ情報のどちらかを統合的に選択し、該当制御シグナルを図4のスイッチングデバイス430に与える。

【0070】

図11のフロー図1100は、各タグスイッチコントローラ1031～1034によって実行される処理を示している。例示のタグスイッチコントローラ1031を使用して、回線交換コントローラ420からの入力と高速メモリ1021からの出力はモニタされ、これは処理ブロック1110で行われる。高速メモリ1021から受信された入力なければ、着信パケットは、回線交換コントローラ420で回線交換される。判定ブロック1120は、高速メモリ1021からの入力があるかどうかを判断するために使用される。入力があれば、処理ブロック1130が起動されるので、タグスイッチコントローラ1031は、スイッチングデバイス430に要求されるステートを、高速メモリの入力から判断することができる。次に、処理ブロック1160が起動され、スイッチングデバイス430を制御するための制御シグナルがタグスイッチコントローラ1031から送信される。高速メモリからの入力がないければ、判定ブロック1140が起動され、回線交換コントローラ420からの入力があるかどうか判断される。回線交換コントローラ420からの入力があれば、ブロック1150による処理が行われ

、タグスイッチコントローラ1031は、スイッチングデバイス430に要求されるステートを、回線交換コントローラ420の入力から判断する。処理ブロック1160は、処理ブロック1150の結果によって再び起動される。回線交換コントローラ1140からの入力がないか、処理ブロック1160が完了すると、コントロールは処理ブロック1110に戻される。

【0071】

繰り返して言えば、光タグスイッチングによれば、高ボリュームバースト、低ボリュームバースト、回線交換トラフィックといった、あらゆるタイプのトラフィックが処理される。これは、タグスイッチ・ネットワークコントロールの2層プロトコルが相互に作用し合うことによって行われる。従って、分散スイッチングコントロールは、シグナリングヘッダを高速に検出し、パケットを該当デスティネーションにルーティングする。長いストリームのパケットが、同じデスティネーションをもつネットワークエレメントに到着したときは、分散スイッチングコントロールは、フロースイッチングコネクションを確立し、パケットストリーム全体が新しく確立したコネクションを通るように転送される。

【0072】

タグスイッチング方法は、波長の数およびノードの数と共に拡張される。これは、分散ノードがマルチ波長シグナリング情報を並行に処理し、これらのノードには、予測スイッチング遅延が、フィルタディレイラインの形で組み込まれていることによるものである。さらに、タグスイッチングは、競合の解決のために経路偏向と波長変換を利用している。

【0073】

(光技術)

光テクノロジーは、本発明を実現するいくつかの重要な側面をカバーしている。この中には光ヘッダ技術、光多重伝送技術、光スイッチング技術、および波長変換技術が含まれている。

【0074】

(a) 光ヘッダ技術

光ヘッダ技術には、図3および図3を参照して上述した光ヘッダ符号化と、光

ヘッダ除去が含まれている。実際には、光ヘッダ210は、ネットワークエレメントへのシグナリングメッセンジャーの働きをし、パケットのデスティネーション、ソース、および長さをネットワークエレメントに通知する。ヘッダ210は、実際のデータペイロードに対して時間的に変位している。これにより、データペイロードは、任意のデータレート／プロトコルまたはフォーマットをもつことが可能になっている。

【0075】

図7と図8を参照して上述したように、ヘッダ符号化は、搬送波をベースにしている。この方法によると、ヘッダ210を変調周波数で分離できるので、ヘッダ検出が単純化される。ヘッダ210は、時間ドメインでデータペイロードの前に置かれているので、最高データレートよりも周波数搬送波が高くなっている。この結果、ヘッダ210の読み取りが可能になり、最終的には、データペイロードに影響することなくヘッダ210の除去が可能になる。

【0076】

(b) 光多重伝送技術

光多重伝送は、図示の例のように、シリカ・アレイ導波管格子構造 (silica arrayed waveguide grating structure) を使用して実装することができる。この導波管格子構造には、次のような独特の利点がある。すなわち、低コスト、スケーラビリティ、低損失、均一性およびコンパクト性である。

【0077】

(c) 光スイッチング技術

高速光スイッチは、非常に長いファイバ遅延を、バッファとして必要としないでパケットルーティングを達成するのに不可欠である。

【0078】

マイクロマシン・エレクトロメカニカルスイッチ (Micromachined Electro Mechanical Switch) によると、スケーラビリティ、低損失、極性無感度、高速スイッチング、および無故障動作といった、望ましい特性を最良に組み合わせたものが得られる。MEMベースの光追加・ドロップスイッチに関して最近報告された結果によれば、9マイクロ秒のスイッチング時間が達成されている。

【0079】

(d) 波長変換技術

波長変換によると、経路偏向またはパケットバッファリングを不要にしてパケットコンテンションが解決される。経路偏向とパケットバッファリングのどちらにも、一連パケットのシーケンスを乱す危険がある。さらに、パケットバッファリングは持続時間だけでなく、容量にも制約があり、非透過方法が要求されることが多い。他方、波長変換は、同一経路を通るように代替波長で伝送することによってブロッキングを解決するので、同一遅延が得られることになる。図示の例では、制限波長変換能力を持つWSXCが配置されている。

【0080】

以上、本発明の教示事項を取り入れた種々実施形態を示し、詳細に説明してきたが、この分野の精通者ならば容易に理解されるように、上述した教示事項を取り入れた多数の他の実施形態を実現することが可能である。

【図面の簡単な説明】

以下では、本発明の教示事項の理解を容易にするために、添付図面を参照して詳細に説明することにする。

【図1】

本発明の一側面によって実施されている、ネットワークの光層と電気層間の結合を示す一般的ネットワークを図示したものである。

【図2】

図1のネットワークの光層を示す図であり、光シグナルヘッダとデータペイロードの相互関係と、ネットワークセッアップでヘッダ/ペイロードがどのように使用されるかを示したものである。

【図3】

ヘッダ符号化とヘッダ除去のための本発明による1つのプラグアンドプレイモジュールを示すハイレベルブロック図である。

【図4】

パケットがWDMネットワークエレメントを通るルートを選択するための本発明による別のプラグアンドプレイモジュールを示すハイレベルブロック図である。

【図5】

WDM回線交換バックボーンネットワークを示す図である。

【図6】

組み込みスイッチを持つ図1のネットワークエレメントと、ローカル・ルーティングテーブルの使用を示す図である。

【図7】

図3に示すプラグアンドプレイモジュール用のヘッダエンコーダ回路の実施例を示すブロック図である。

【図8】

図3に示すプラグアンドプレイモジュール用のヘッダリムーバ回路の実施例を示すブロック図である。

【図9】

図4に示すプラグアンドプレイモジュール用のヘッダ検出器回路の実施例を示すブロック図である。

【図10】

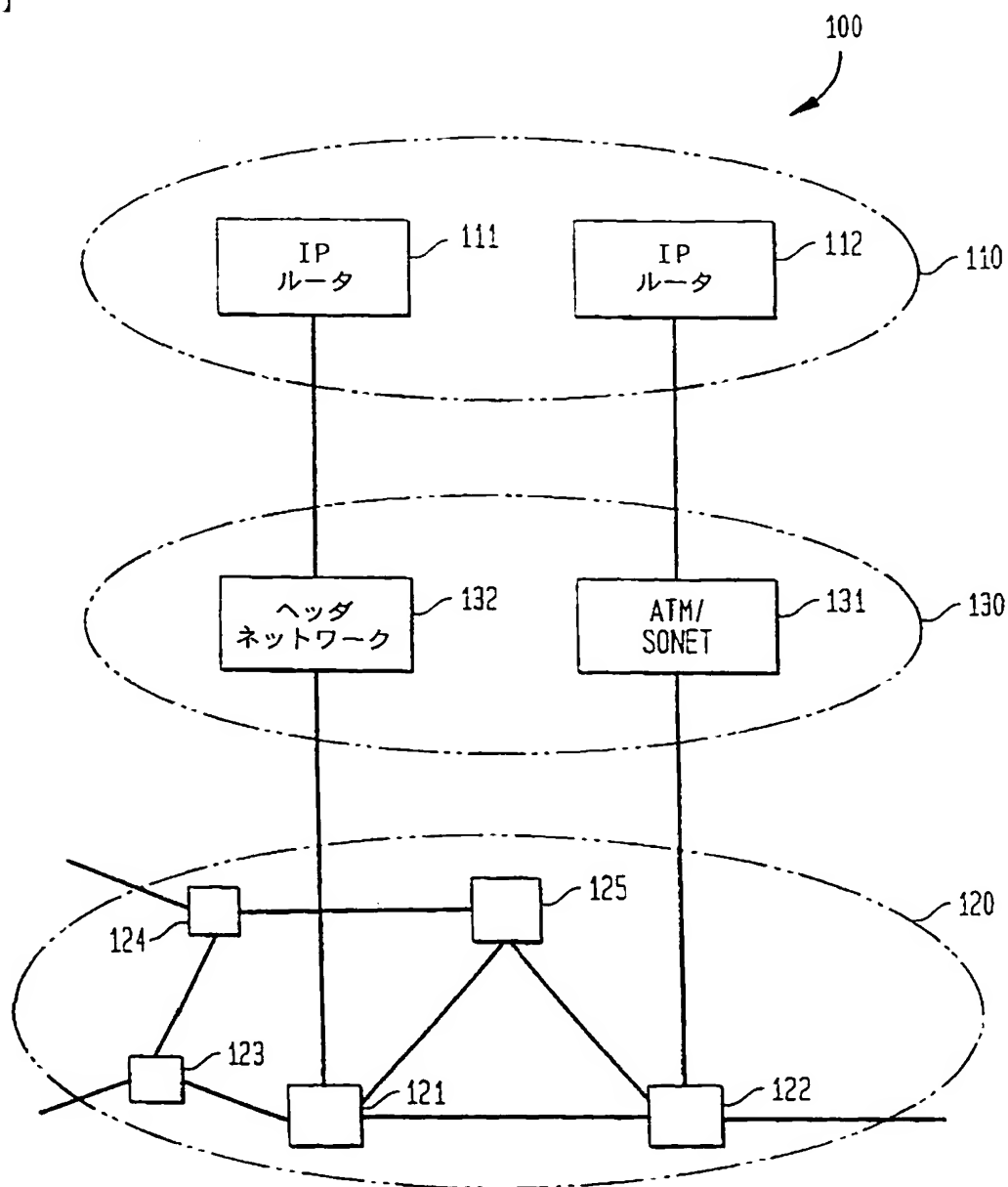
図4の実施例の詳細を示すブロック図であり、タグスイッチコントローラは、中間に置かれたデマルチプレクサ、およびヘッダ検出器と高速メモリを備えていることが示されている。

【図11】

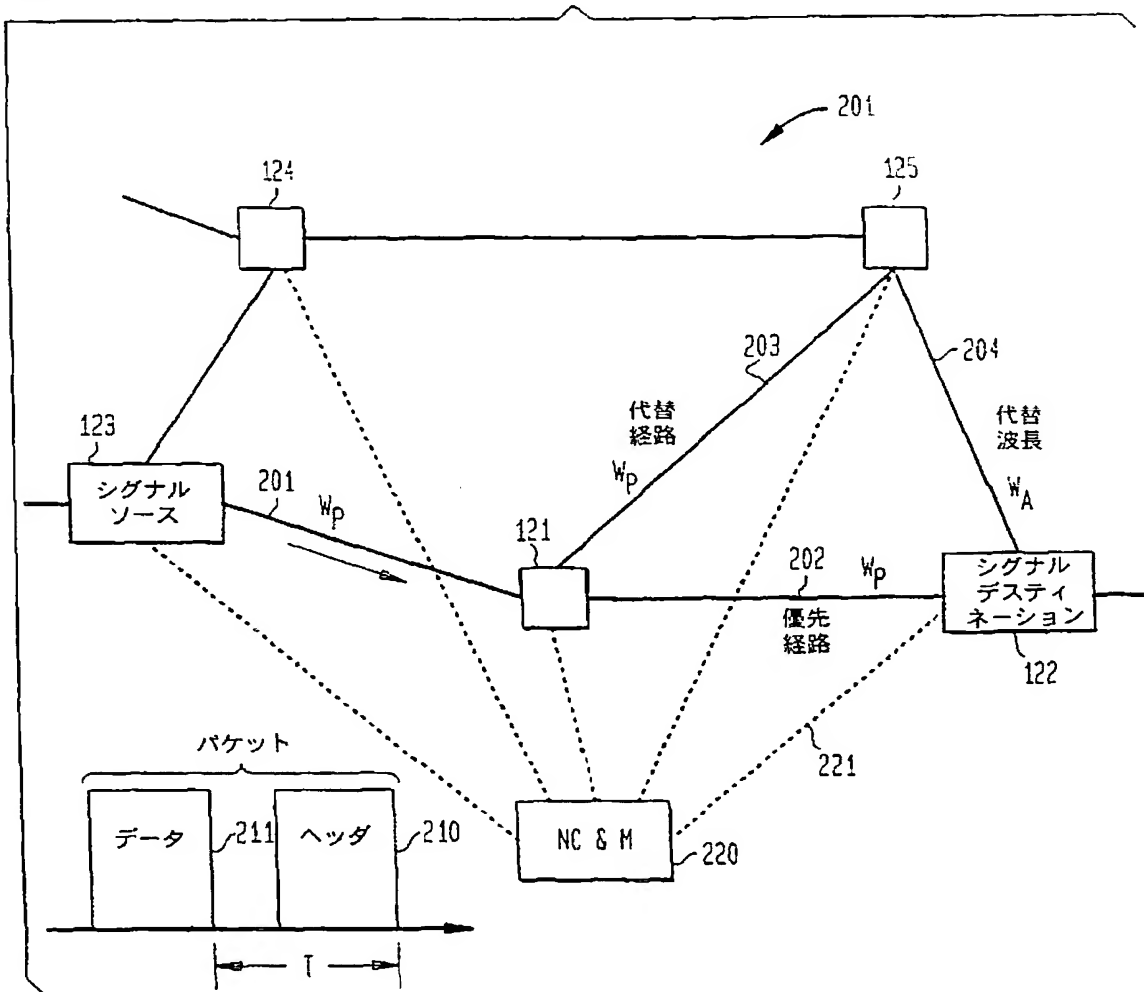
図10の各タグスイッチコントローラによって実行される処理を示すフロー図である。

添付図面において、理解を容易にするために、各図に共通する同一エレメントは、可能な限り、同一参照符号を使用して示されている。

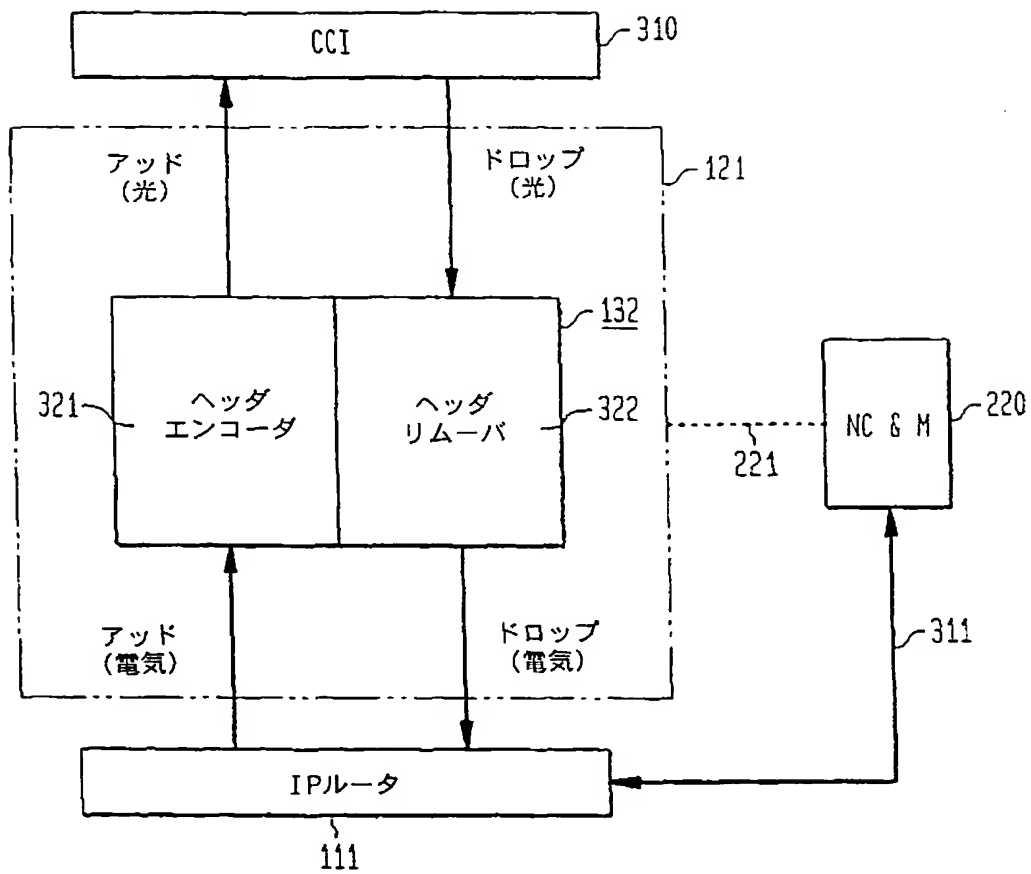
【図1】



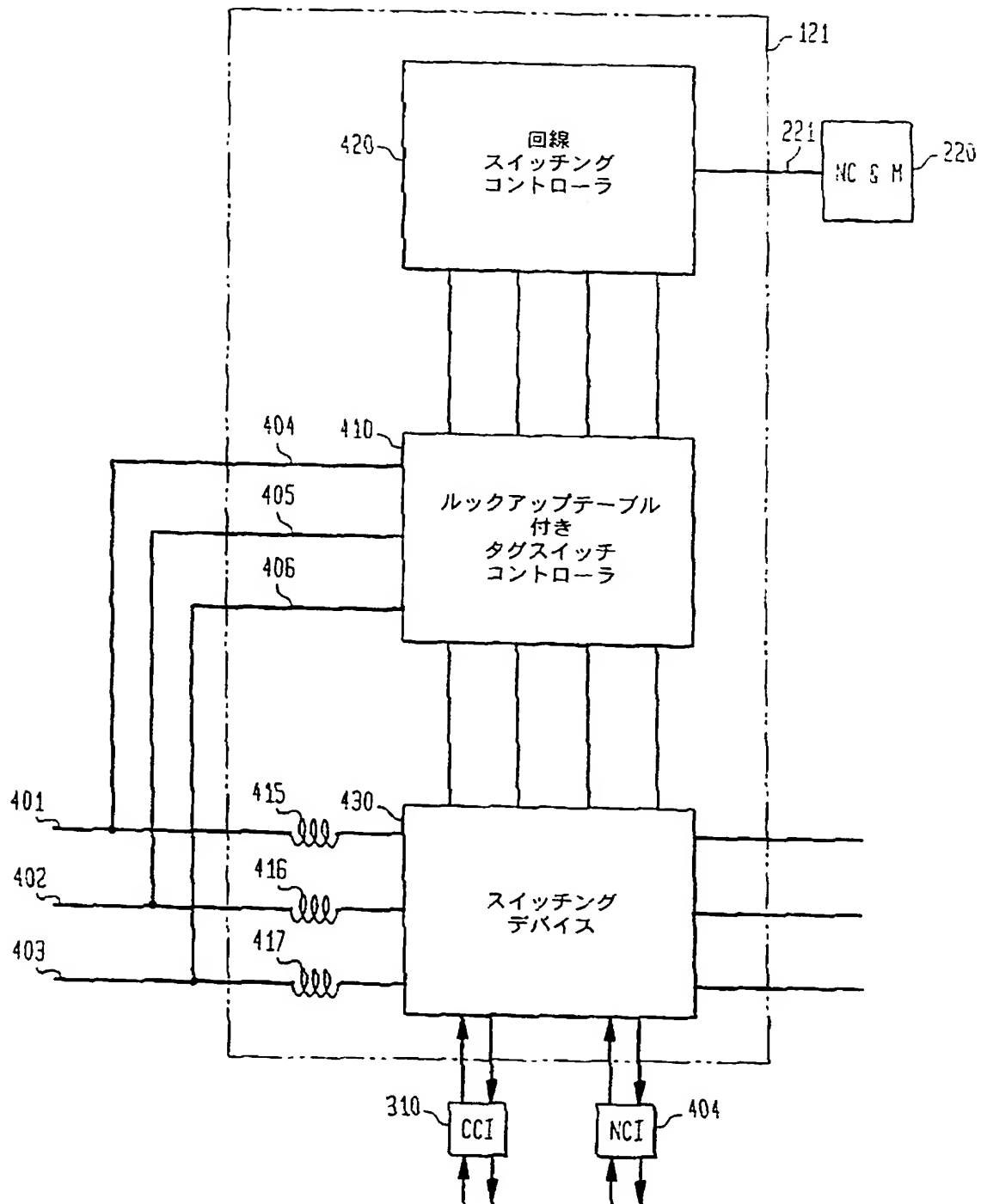
【図2】



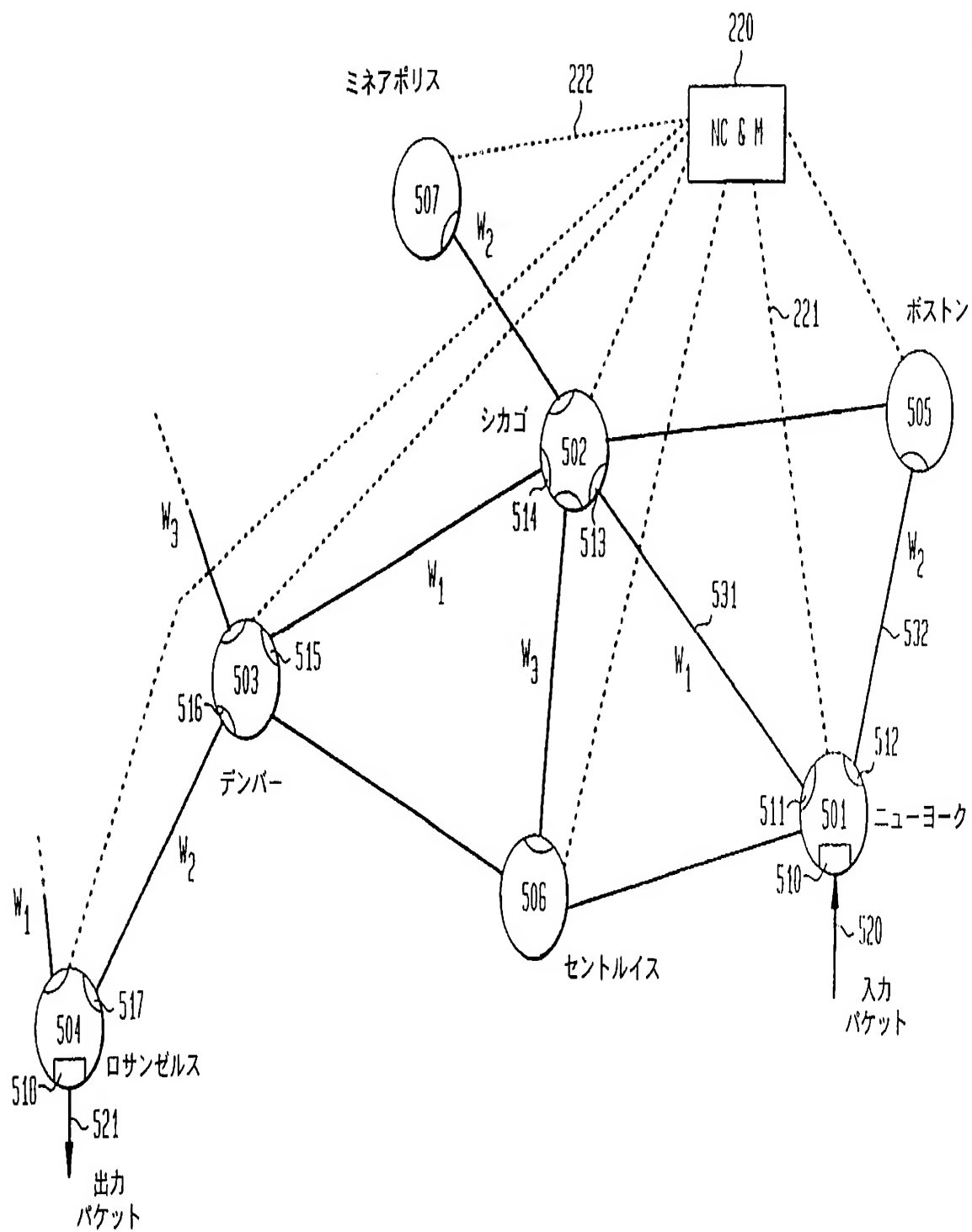
【図3】



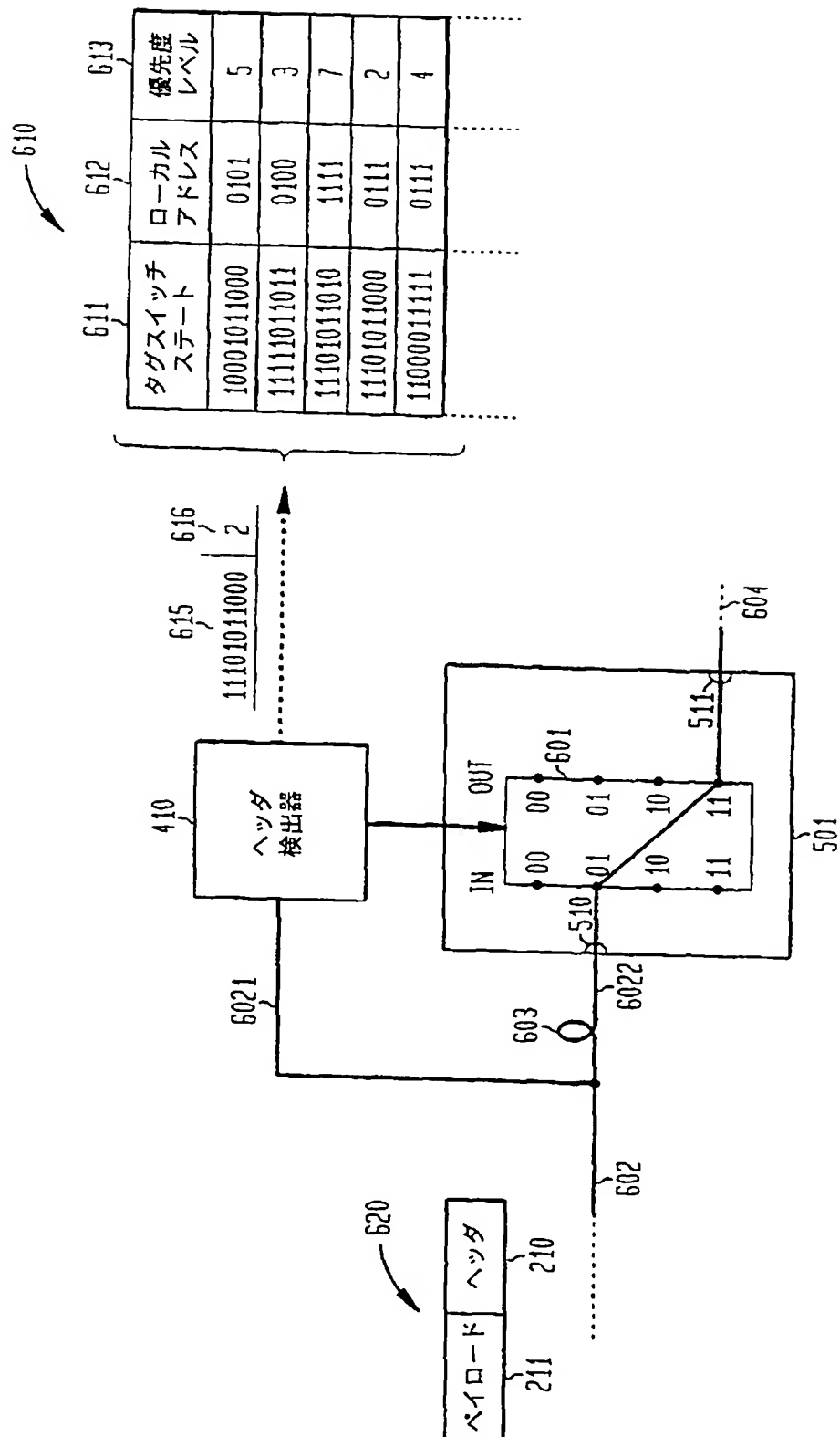
【図4】



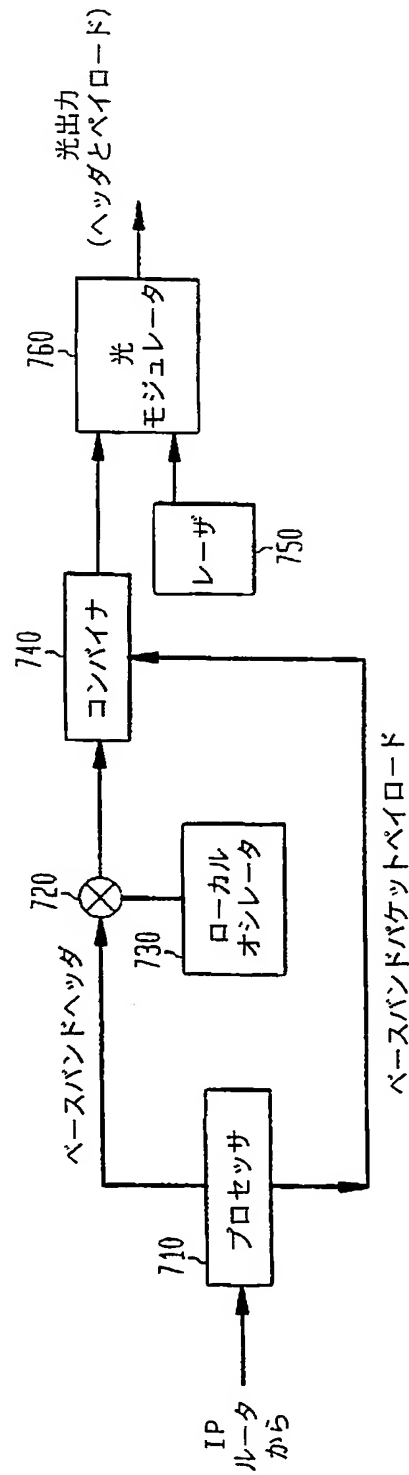
【図 5】



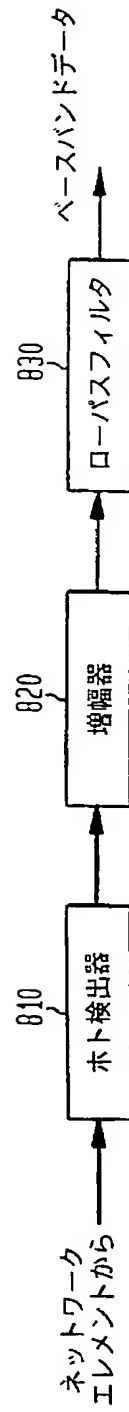
【図6】



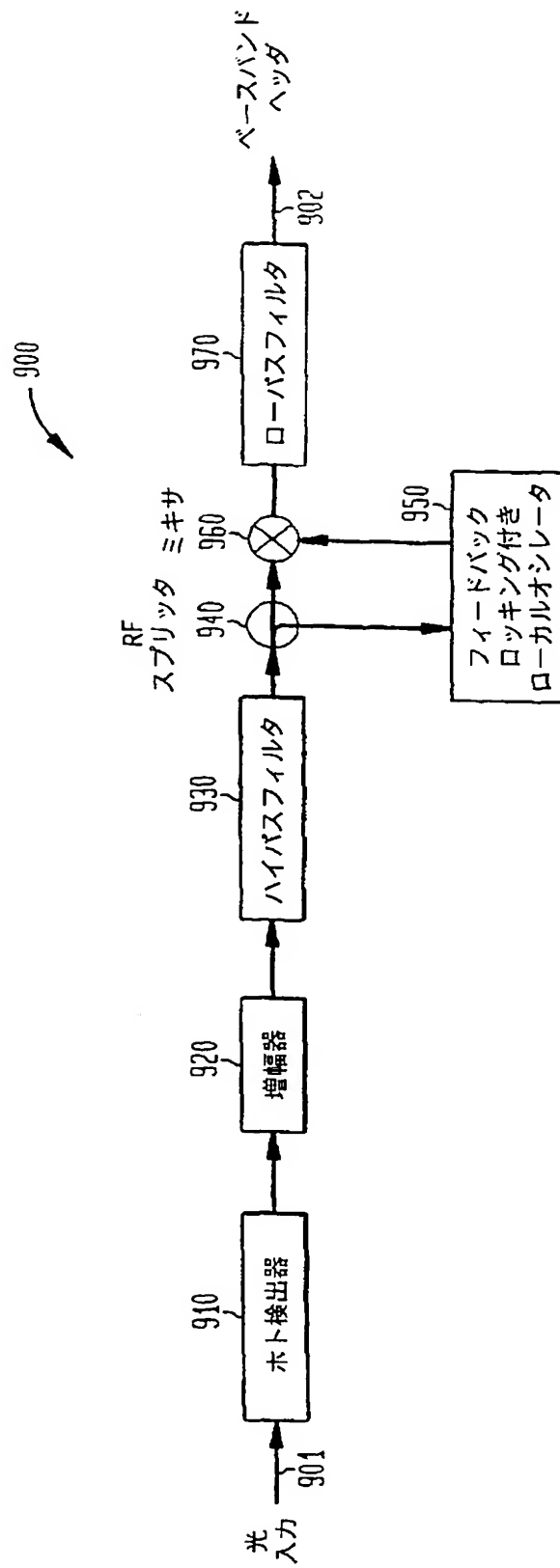
【図7】

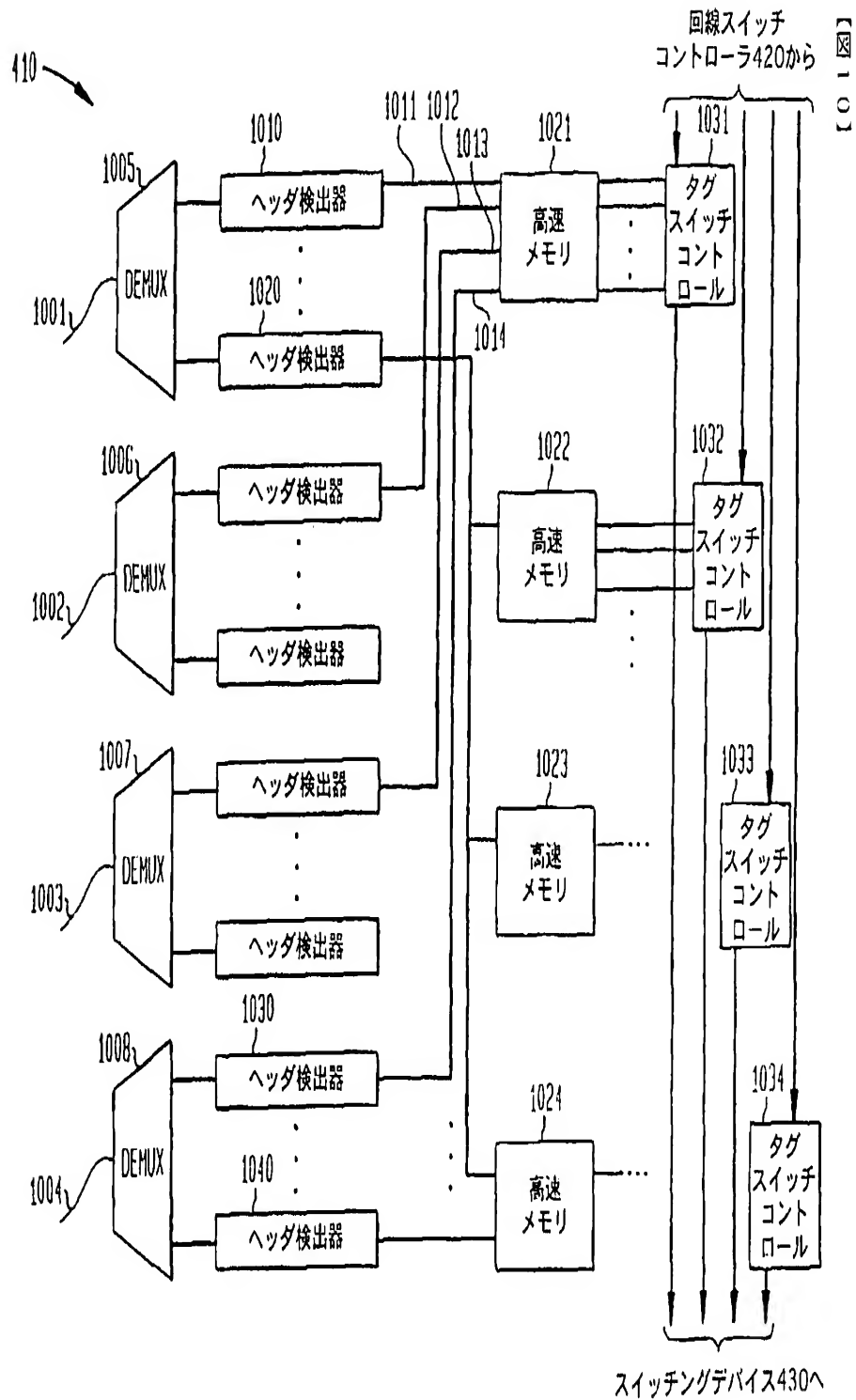


【図8】

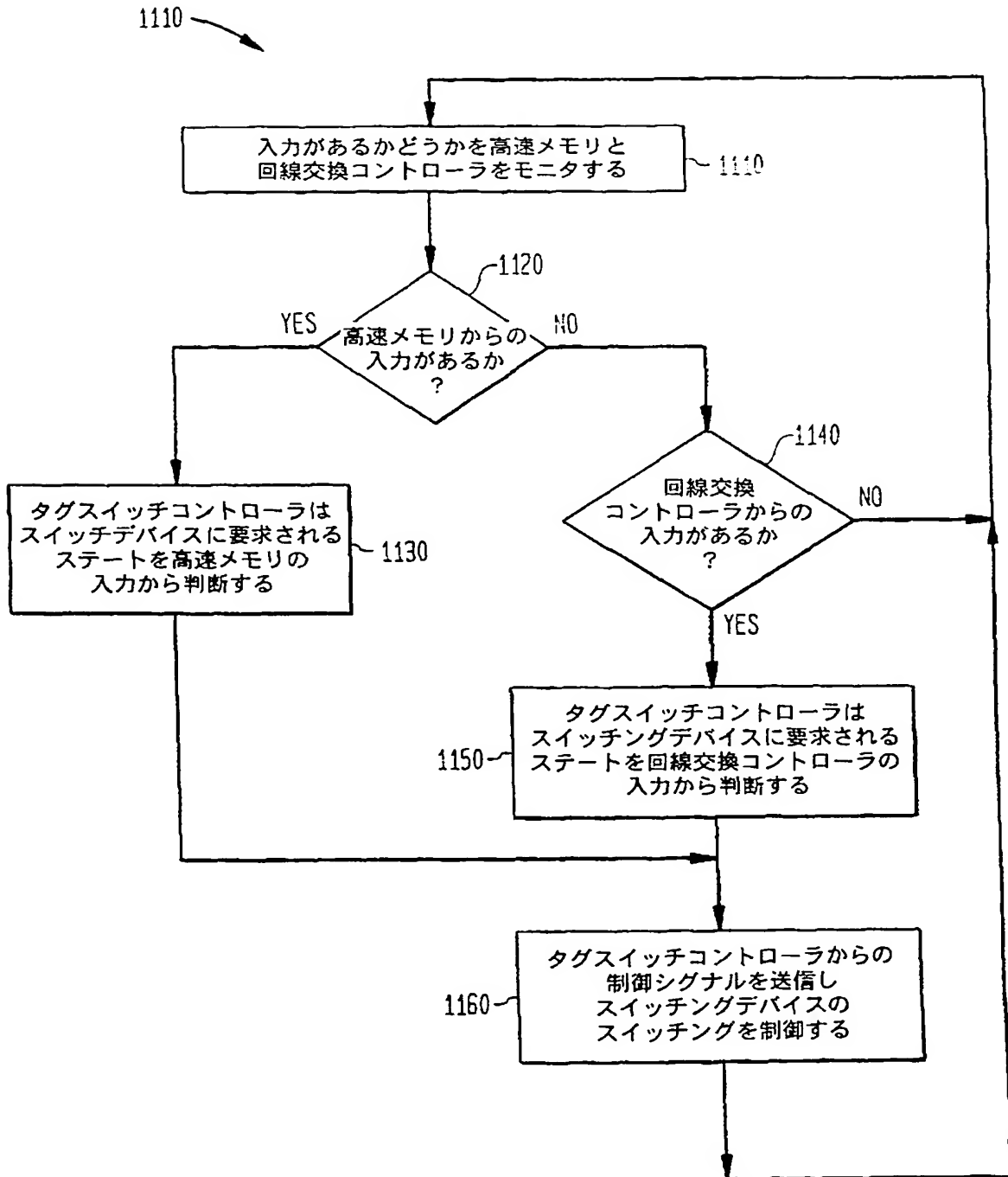


【図9】





【図11】



【手続補正書】

【提出日】平成13年4月19日（2001. 4. 19）

【手続補正1】

【補正対象書類名】明細書

【補正対象項目名】特許請求の範囲

【補正方法】変更

【補正内容】

【特許請求の範囲】

【請求項1】 波長分割多重伝送（WDM）ネットワーク内で入力ネットワークエレメントから出力ネットワークエレメントへデータペイロードを伝達するための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有する方法において、

ローカル・ルックアップテーブルを生成し、前記ネットワークエレメントの各々にストアするステップであって、各々のローカル・ルックアップテーブルは、関連する前記ネットワークエレメントの1つを通るローカルルートを決するステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に前記データペイロードと同じ波長に埋め込まれた前記データペイロードに光ヘッダを追加するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードとヘッダがWDMネットワーク内を伝達するときに前記ネットワークエレメントで前記ヘッダを光学的に判断するステップと、

前記対応するローカル・ルックアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記ネットワークエレメントを通る前

記データペイロードと前記ヘッダのルート選択をするステップと
を備えたことを特徴とする方法。

【請求項2】 前記光ヘッダは、前記ネットワークエレメントを通る前記光ヘッダと前記データペイロードのルート選択のためのタグスイッチステートを含み、光ヘッダを追加するステップは、前記入力ネットワークエレメントから前記出力ネットワークエレメントへ前記ネットワークエレメントを通る前記光ヘッダと前記データペイロードのルート選択のための該当するタグスイッチステートを判断し、前記光ヘッダに挿入するステップを含むことを特徴とする請求項1に記載の方法。

【請求項3】 前記光ヘッダを前記データペイロードに追加するステップは、時間的に前記データペイロードより進んだ位置に前記光ヘッダを置くステップを含むことを特徴とする請求項1に記載の方法。

【請求項4】 ルート選択をするステップは、前記選択されたルートの競合を解決するステップを含むことを特徴とする請求項1に記載の方法。

【請求項5】 競合を解決するステップは、前記選択されたルートに関連して決定された代替ルートをルート選択するステップを含むことを特徴とする請求項4に記載の方法。

【請求項6】 競合を解決するステップは、前記選択されたルートに使用された波長に関連して決定された代替波長をルート選択するステップを含むことを特徴とする請求項4に記載の方法。

【請求項7】 前記ヘッダを光学的に判断するステップは、前記データペイロードが妨げられない光ドメインに残っている間に、前記ネットワークエレメントで前記ヘッダを光電磁気学的に判断することを特徴とする請求項1に記載の方法。

【請求項8】 複数のネットワークエレメントで構成された波長分割多重伝送(WDM)ネットワーク内で入力ネットワークエレメントから出力ネットワークエレメントへデータペイロードのシーケンスを伝達するための方法であって、前記データペイロードの各々は、所定のフォーマットとプロトコルを有する方法において、

ローカル・ルックアップテーブルを生成し、前記ネットワークエレメントの各々にストアするステップであって、各々のローカル・ルックアップテーブルは、関連する前記ネットワークエレメントの1つを通るローカルルートを決定するステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に前記データペイロードと同じ波長に埋め込まれた前記データペイロードの各々に光ヘッダを追加するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードの各々とそれぞれに対応するヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを示しており、前記データペイロードの各々の前記フォーマットとプロトコルは、それぞれに対応するヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードの各々とそれぞれに対応するヘッダがWDMネットワーク内を伝達するときに前記ネットワークエレメントで前記ヘッダを光学的に判断するステップと、

前記対応するローカル・ルックアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードの第1とそれに対応するヘッダが前記ネットワークエレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記ネットワークエレメントを通る前記データペイロードの第1とそれに対応するヘッダのルート選択をするステップと、

前記データペイロードの第1で選択された前記ローカルルートを通して前記シーケンス中の前記データペイロードの後続する各々のルート選択をするステップと

を備えたことを特徴とする方法。

【請求項9】 ルート選択する各ステップは、前記選択されたルートの競合を解決するステップを含むことを特徴とする請求項8に記載の方法。

【請求項10】 競合を解決するステップは、前記選択されたルートに関連して決定された代替ルートをルート選択するステップを含むことを特徴とする請

求項9に記載の方法。

【請求項11】 競合を解決するステップは、前記選択されたルートに使用された波長に関連して決定された代替波長をルート選択するステップを含むことを特徴とする請求項9に記載の方法。

【請求項12】 入力ネットワークエレメントに到着したデータペイロードを、複数のネットワークエレメントで構成された波長分割多重伝送（WDM）ネットワーク上を伝達させるための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有する方法において、

前記データペイロードに関連付けられた光ヘッダを生成するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に前記データペイロードと同じ波長に埋め込まれた前記データペイロードに前記光ヘッダを追加するステップと

を備えたことを特徴とする方法。

【請求項13】 複数のネットワークエレメントで構成された波長分割多重伝送（WDM）ネットワーク内で各々の特定のネットワークエレメントの入力から出力にヘッダとデータペイロードを転送するための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有し、ヘッダのフォーマットとプロトコルから独立している方法において、

ローカル・ルックアップテーブルを生成し、前記特定のネットワークエレメントにストアするステップであって、前記ローカル・ルックアップテーブルは、前記特定のネットワークエレメントを通るローカルルートを決めるステップと、

前記ペイロードとヘッダが前記特定のネットワークエレメントの前記入力に到達したときに前記ヘッダを光学的に判断するステップと、

前記ローカル・ルックアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが前記特定のネットワーク

エレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記特定のネットワークエレメントを通る前記データペイロードと前記ヘッダのルート選択をするステップと

を備えたことを特徴とする方法。

【請求項14】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される （a） 電気層と （b） 光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するためのシステムであって、前記データペイロードは、所定のフォーマットとプロトコルを有するシステムにおいて、

前記ソースデバイスと前記WDMネットワークを結合し、前記データペイロードを前記WDMネットワークに入力する前に光ヘッダを前記データペイロードの前に追加するための光ヘッダモジュールの第1のタイプであって、前記ヘッダは、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している光ヘッダモジュールの第1のタイプと、

各々の前記ネットワークエレメントに付加された光ヘッダモジュールの第2のタイプであって、ローカル・ルックアップテーブルに対応する前記ネットワークエレメントの1つにストアする手段であって、各々のローカル・ルックアップテーブルは、対応する前記ネットワークエレメントの1つを通るルーティング経路を決定する手段と、前記データペイロードとヘッダがWDMネットワーク上を伝達するときに対応する前記ネットワークエレメントの1つで前記ヘッダを光学的に判断する手段と、前記対応するローカル・ルックアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが対応する前記ネットワークエレメントの1つを通る前記ローカルルートを選択する手段と、前記選択されたルートに対応するように対応する前記ネットワークエレメントの1つを通る前記データペイロードと前記ヘッダのルート選択をする手段とを含む光ヘッダモジュールの第2のタイプと

を備えたことを特徴とするシステム。

【請求項15】 前記光ヘッダモジュールの第1のタイプ他方は、WDMネットワークをデスティネーションデバイスに結合しており、前記光ヘッダモジュールの第1のタイプは、前記デスティネーションデバイスに配達する前に前記データペイロードから前記ヘッダを除去する手段を含むことを特徴とする請求項14に記載のシステム。

【請求項16】 前記光ヘッダモジュールの第2のタイプは、前記データペイロードが妨げられない光ドメインに残っている間に、前記ヘッダを光電磁気学的に判断する手段をさらに含むことを特徴とする請求項14に記載のシステム。

【請求項17】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される（a）電気層と（b）光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するための光ヘッダモジュールであって、前記データペイロードは、所定のフォーマットとプロトコルを有する光ヘッダモジュールにおいて、前記光ヘッダモジュールは、前記ソースデバイスと前記WDMネットワークを結合し、前記光ヘッダモジュールは、前記データペイロードに関連付けられた光ヘッダを生成する手段であって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している手段と、前記データペイロードを前記入力ネットワークエレメントに入力する前に前記光ヘッダを前記データペイロードに追加する手段であって、前記ペイロードと同じ波長の前記光ヘッダに埋め込む手段を含む前記追加する手段とを含むことを特徴とする光ヘッダモジュール。

【請求項18】 前記光ヘッダは、前記ネットワークエレメントを通る前記光ヘッダと前記データペイロードのルート選択のためのタグスイッチステートを含み、光ヘッダを追加する手段は、前記ネットワークエレメントを通る前記光ヘッダと前記データペイロードのルート選択のための該当するタグスイッチステートを判断し、前記光ヘッダに挿入する手段を含むことを特徴とする請求項17に

記載のモジュール。

【請求項19】 前記光ヘッダを前記データペイロードに追加する手段は、時間的に前記データペイロードより進んだ位置に前記光ヘッダを置く手段を含むことを特徴とする請求項17に記載のモジュール。

【請求項20】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される（a）電気層と（b）光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイス向けに送られるデータペイロードを伝達するための光ヘッダプロセッサであって、前記データペイロードは、所定のフォーマットとプロトコルを有する光ヘッダプロセッサにおいて、前記データペイロードは、さらに関連付けられた光ヘッダを有し、前記光ヘッダプロセッサは、各々の前記ネットワークエレメントと関連付けられ、前記光ヘッダプロセッサは、

ローカル・ルップアップテーブルを生成し、対応する前記ネットワークエレメントの1つにストアする手段であって、各々のローカル・ルップアップテーブルは、対応する前記ネットワークエレメントの1つを通るルーティング経路を決定する手段と、

前記データペイロードとヘッダがWDMネットワーク内を伝達するときに対応する前記ネットワークエレメントの1つで前記ヘッダを光学的に判断する手段と、

前記対応するローカル・ルップアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが対応する前記ネットワークエレメントの1つを通る前記ローカルルートを選択する手段と、

前記選択されたルートに対応するように対応する前記ネットワークエレメントの1つを通る前記データペイロードと前記ヘッダのルート選択をする手段と

を備えたことを特徴とする光ヘッダプロセッサ。

【請求項21】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される（a）電気層と（b）光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するためのシステムで

あって、前記データペイロードは、所定のフォーマットとプロトコルを有し、前記ネットワークは、前記ネットワークエレメントに結合され、前記ネットワークを通る回線交換ルートを決定するネットワークマネージャを含み、各々の前記ネットワークエレメントは、(i) スイッチングデバイスと、(ii) 前記ネットワークマネージャに応答して、前記ネットワークマネージャからの入力に基づいて前記スイッチングデバイスを制御して、前記WDMネットワークを通る回線交換ルーティング経路を確立する回線交換コントローラを含むシステムにおいて、

前記ソースデバイスと前記WDMネットワークを結合し、前記データペイロードを前記WDMネットワークに入力する前に光ヘッダを前記データペイロードの前に追加するための光ヘッダモジュールの第1のタイプであって、前記ヘッダは、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している光ヘッダモジュールの第1のタイプと、

前記ネットワークマネージャと前記回線交換コントローラに응答し、前記スイッチングデバイスに結合されている光ヘッダモジュールの第2のタイプであって、前記ネットワークマネージャから与えられたローカル・ルップアップテーブルを各々のネットワークエレメントにストアする手段であって、各々のローカル・ルップアップテーブルは、各々のネットワークエレメントを通るルーティング経路を決定する手段と、前記データペイロードと前記ヘッダが前記WDMネットワーク内を伝達されるときに各々の前記ネットワークエレメントで前記ヘッダを光学的に判断する手段と、前記対応するローカル・ルップアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを選択する手段と、前記回線交換コントローラからの入力と前記ローカル・ルップアップテーブルを処理して前記スイッチングデバイスを制御することによって、前記選択されたルートに対応するように各々の前記ネットワークを通る前記データペイロードと前記ヘッダのルート選択をする手段と

を備えたことを特徴とするシステム。

【請求項22】 前記スイッチングデバイスの前に置かれていて、前記スイッチングデバイスへの前記データペイロードと前記ヘッダの配達を、あらかじめ決めたインターバルで遅延させる手段をさらに含むことを特徴とする請求項21に記載のシステム。

【請求項23】 各々の前記ネットワークエレメントで前記ヘッダを光学的に判断する手段は、前記WDMネットワーク上を伝達する前記ヘッダをベースバンドのヘッダに復調するデマルチプレクサを含み、

選択する手段は、前記デマルチプレクサに応答して、前記ベースバンドのヘッダに含まれるルート情報を判断するための高速メモリを含み、

ルートを選択する手段は、前記ローカル・ルックアップテーブルに結合され、前記高速メモリと前記回線交換コントローラにตอบสนองして、前記スイッチングデバイスを制御するためのタグスイッチコントローラを含むことを特徴とする請求項21に記載のシステム。

【請求項24】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される（a）電気層と（b）光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイスへ送られるデータペイロードを伝達するためのシステムであって、前記データペイロードは、所定のフォーマットとプロトコルを有するシステムにおいて、

前記ソースデバイスと前記WDMネットワークを結合し、前記データペイロードを前記WDMネットワークに入力する前に光ヘッダを前記データペイロードの前に追加するための光ヘッダモジュールの第1のタイプであって、前記ヘッダは、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している光ヘッダモジュールの第1のタイプと、

各々の前記ネットワークエレメントに付加された光ヘッダモジュールの第2のタイプであって、ローカル・ルックアップテーブルを対応する前記ネットワーク

エレメントの1つにストアする手段であって、各々のローカル・ルックアップテーブルは、対応する各々の前記ネットワークエレメントの1つを通るルーティング経路を決定する手段と、前記データペイロードとヘッダが前記WDMネットワーク上を伝達するときに対応する前記ネットワークエレメントの1つで前記ヘッダを光学的に判断する手段と、前記対応するローカル・ルックアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが対応する前記ネットワークエレメントの1つを通るローカルルートを選択する手段と、前記選択されたルートに対応するように対応する前記ネットワークエレメントの1つを通る前記データペイロードと前記ヘッダのルート選択をする手段と、同じローカルルートを有する後続の連続するヘッダの各々に対して前記選択されたルートを維持する手段とを含む光ヘッダモジュールの第2のタイプと

を備えたことを特徴とするシステム。

【請求項25】 波長分割多重伝送（WDM）ネットワーク内で入力ネットワークエレメントから出力ネットワークエレメントへデータペイロードを伝達するための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有する方法において、

ローカル・ルックアップテーブルを生成し、前記ネットワークエレメントの各々にストアするステップであって、各々のローカル・ルックアップテーブルは、関連する前記ネットワークエレメントの1つを通るローカルルートを決めるステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に光ヘッダを前記データペイロードに追加するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードとヘッダがWDMネットワーク内を伝達するときに前記ネットワークエレメントで前記ヘッダを光学的に判断するステップと、

前記対応するローカル・ルックアップテーブル内の前記ヘッダを調べることに
よって決定された通りに、前記データペイロードと前記ヘッダが前記ネットワー
クエレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記ネットワークエレメントを通る前
記データペイロードと前記ヘッダのルート選択をするステップとを含み、

前記光ヘッダは、前記ネットワークエレメントを通る前記光ヘッダと前記デー
タペイロードのルート選択のためのタグスイッチステートを含み、光ヘッダを追
加するステップは、前記入力ネットワークエレメントから前記出力ネットワー
クエレメントへ前記ネットワークエレメントを通る前記光ヘッダと前記データペ
イロードのルート選択のための該当するタグスイッチステートを判断し、前記光ヘ
ッダに挿入するステップを含み、

前記光ヘッダは、前記光ヘッダと前記データペイロードが前記ネットワークエ
レメントを通過して伝達するときにルート競合を解決する際に使用する優先度デー
タをさらに含み、ローカル・ルックアップテーブルを生成し、ストアするステッ
プは、優先度レベルを各々のタグスイッチステートに関連付けるステップを含み
、前記光ヘッダを追加するステップは、前記データペイロードの該当する優先度
データを前記光ヘッダに挿入するステップを含み、選択するステップは、前記優
先度データと前記優先度レベルに基づいて前記ローカルルートを判断するステッ
プを含むことを特徴とする方法。

【請求項26】 波長分割多重伝送(WDM)ネットワーク内で入力ネット
ワークエレメントから出力ネットワークエレメントへデータペイロードを伝達す
るための方法であって、前記データペイロードは、所定のフォーマットとプロト
コルを有する方法において、

ローカル・ルックアップテーブルを生成し、前記ネットワークエレメントの各
々にストアするステップであって、各々のローカル・ルックアップテーブルは、
関連する前記ネットワークエレメントの1つを通るローカルルートを決めるス
テップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に光ヘ
ッダを前記データペイロードに追加するステップであって、前記ヘッダは、フォ

ーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードとヘッダがWDMネットワーク内を伝達するときに前記ネットワークエレメントで前記ヘッダを光学的に判断するステップと、

前記対応するローカル・ルックアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記ネットワークエレメントを通る前記データペイロードと前記ヘッダのルート選択をするステップとを備え、

前記光ヘッダと前記データペイロードは、最初にベースバンドで生成され、前記光ヘッダを前記データペイロードに追加するステップは、

前記ベースバンド光ヘッダを、前記ベースバンドデータペイロードの周波数帯域より上の周波数帯域に周波数シフトするステップと、

前記周波数シフトされたベースバンド光ヘッダと前記ベースバンドデータペイロードを結合して複合周波数信号を形成するステップと、

所定の波長の光源を使用して前記複合周波数信号を光変調して、WDMネットワークを通して前記ヘッダと前記データペイロードを伝達する光信号を生成するステップと

を含むことを特徴とする方法。

【請求項27】 波長分割多重伝送（WDM）ネットワーク内で入力ネットワークエレメントから出力ネットワークエレメントへデータペイロードを伝達するための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有する方法において、

ローカル・ルックアップテーブルを生成し、前記ネットワークエレメントの各々にストアするステップであって、各々のローカル・ルックアップテーブルは、関連する前記ネットワークエレメントの1つを通るローカルルートを決めるステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に光ヘッダを前記データペイロードに追加するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通る前記ローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードとヘッダがWDMネットワーク内を伝達するときに前記ネットワークエレメントで前記ヘッダを光学的に判断するステップと、

前記対応するローカル・ルックアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが前記ネットワークエレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記ネットワークエレメントを通る前記データペイロードと前記ヘッダのルート選択をするステップとを備え、

前記光ヘッダと前記データペイロードは、最初にベースバンドで生成され、前記光ヘッダを前記データペイロードに追加するステップは、

前記ベースバンド光ヘッダを、前記ベースバンドデータペイロードの周波数帯域より上の周波数帯域に周波数シフトするステップと、

前記周波数シフトされたベースバンド光ヘッダと前記ベースバンドデータペイロードを結合して複合周波数信号を形成するステップと、

所定の波長の光源を使用して前記複合周波数信号を光変調して、WDMネットワークを通して前記ヘッダと前記データペイロードを伝達する光信号を生成するステップとを含み、

各々の前記ネットワークエレメントで前記ヘッダを光学的に判断するステップは、

前記光ヘッダをホト検出して検出信号を生成するステップと、

ローカルロッキングオシレータで前記検出信号をロッキングして、ロック信号を生成するステップと、

前記検出信号と前記ロック信号を混合して、ベースバンドの前記ヘッダを表すベースバンド信号を出力するステップと

を含むことを特徴とする方法。

【請求項28】 入力ネットワークエレメントに到着したデータペイロードを、複数のネットワークエレメントで構成された波長分割多重伝送(WDM)ネットワーク上を伝達させるための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有する方法において、

前記データペイロードに関連付けられた光ヘッダを生成するステップであって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立しているステップと、

前記データペイロードを前記入力ネットワークエレメントに入力する前に前記光ヘッダを前記データペイロードに追加するステップとを備え、

前記光ヘッダと前記ペイロードは、最初にベースバンドで生成され、前記光ヘッダを前記データペイロードに追加するステップは、

前記ベースバンド光ヘッダを、前記ベースバンドデータペイロードの周波数帯域より上の周波数帯域に周波数シフトするステップと、

前記周波数シフトされたベースバンド光ヘッダと前記ベースバンドデータペイロードを結合して複合周波数シグナルを形成するステップと、

所定の波長の光源を使用して前記複合周波数シグナルを光変調して、WDMネットワークを通して前記ヘッダと前記データペイロードを伝達する光信号を生成するステップと

を含むことを特徴とする方法。

【請求項29】 複数のネットワークエレメントで構成された波長分割多重伝送(WDM)ネットワーク内で各々の特定のネットワークエレメントの入力から出力にヘッダとデータペイロードを転送するための方法であって、前記データペイロードは、所定のフォーマットとプロトコルを有し、ヘッダのフォーマットとプロトコルから独立している方法において、

ローカル・ルックアップテーブルを生成し、前記特定のネットワークエレメントにストアするステップであって、前記ローカル・ルックアップテーブルは、前

記特定のネットワークエレメントを通るローカルルートを決定するステップと、

前記ペイロードとヘッダが前記特定のネットワークエレメントの前記入力に到達したときに前記ヘッダを光学的に判断するステップと、

前記ローカル・ルックアップテーブル内の前記ヘッダを調べることによって決定された通りに、前記データペイロードと前記ヘッダが前記特定のネットワークエレメントを通る前記ローカルルートを選択するステップと、

前記選択されたルートに対応するように前記特定のネットワークエレメントを通る前記データペイロードと前記ヘッダのルート選択をするステップとを備え、

各々の前記ネットワークエレメントで前記ヘッダを光学的に判断するステップは、

前記光ヘッダをホト検出して検出信号を出力するステップと、

ローカルロッキングオシレータで前記検出信号をロッキングして、ロック信号を生成するステップと、

前記検出信号と前記ロック信号を混合して、ベースバンドの前記ヘッダを表すベースバンド信号を出力するステップと

を含むことを特徴とする方法。

【請求項30】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される （a） 電気層と （b） 光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデステーションデバイスへ送られるデータペイロードを伝達するための光ヘッダモジュールであって、前記データペイロードは、所定のフォーマットとプロトコルを有する光ヘッダモジュールにおいて、前記光ヘッダモジュールは、前記ソースデバイスと前記WDMネットワークを結合し、前記光ヘッダモジュールは、前記データペイロードに関連付けられた光ヘッダを生成する手段であって、前記ヘッダは、フォーマットとプロトコルを有し、前記データペイロードと前記ヘッダが各々の前記ネットワークエレメントを通るローカルルートを示しており、前記データペイロードの前記フォーマットとプロトコルは、前記ヘッダの前記フォーマットとプロトコルから独立している手段と、前記データペイロードを前記入力ネットワークエレメントに入力する前に前記光ヘッダを前記データペイロードに追

加する手段とを含み、

前記光ヘッダを前記データペイロードに追加する手段は、時間的に前記データペイロードより進んだ位置に前記光ヘッダを置く手段を含み、

前記光ヘッダと前記データペイロードは、最初にベースバンドで生成され、前記光ヘッダを前記データペイロードに追加する手段は、

前記ベースバンド光ヘッダを、前記ベースバンドデータペイロードの周波数帯域より上の周波数帯域に周波数シフトする手段と、

前記周波数シフトされたベースバンド光ヘッダと前記ベースバンドデータペイロードを結合して複合周波数信号を形成する手段と、

所定の波長の光源を使用して前記複合周波数信号を光変調して、WDMネットワークを通して前記ヘッダと前記データペイロードを伝達する光信号を生成する手段と

を含むことを特徴とする光ヘッダモジュール。

【請求項31】 複数のネットワークエレメントを含む波長分割多重伝送（WDM）ネットワークで構成される （a） 電気層と （b） 光層の組み合わせ構造において、電気層のソースデバイスによって生成され、電気層のデスティネーションデバイス向けに送られるデータペイロードを伝達するための光ヘッダプロセッサであって、前記データペイロードは、所定のフォーマットとプロトコルを有する光ヘッダプロセッサにおいて、前記データペイロードは、さらに関連付けられた光ヘッダを有し、前記光ヘッダプロセッサは、各々の前記ネットワークエレメントと関連付けられ、前記光ヘッダプロセッサは、

ローカル・ルップアップテーブルを生成し、対応する前記ネットワークエレメントの1つにストアする手段であって、各々のローカル・ルップアップテーブルは、対応する前記ネットワークエレメントの1つを通るルーティング経路を決定する手段と、

前記データペイロードとヘッダがWDMネットワーク内を伝達するときに対応する前記ネットワークエレメントの1つで前記ヘッダを光学的に判断する手段と、

前記対応するローカル・ルップアップテーブル内の前記ヘッダを調べることに

よって決定された通りに、前記データペイロードと前記ヘッダが対応する前記ネットワークエレメントの1つを通る前記ローカルルートを選択する手段と、

前記選択されたルートに対応するように対応する前記ネットワークエレメントの1つを通る前記データペイロードと前記ヘッダのルート選択をする手段とを備え、

各々の前記ネットワークエレメントで前記ヘッダを光学的に判断する手段は、

前記光ヘッダをホト検出して検出信号を出力する手段と、

前記検出信号をロッキングしてロック信号を生成するローカルロッキングオシレータと、

前記検出信号と前記ロック信号を混合して、ベースバンドの前記ヘッダを表すベースバンド信号を出力する手段と

を含むことを特徴とする光ヘッダプロセッサ。

【手続補正2】

【補正対象書類名】明細書

【補正対象項目名】0022

【補正方法】変更

【補正内容】

【0022】

図1は、中間層130で光層120と電気層110を結合することによって得られる、汎用ネットワーク100の光層120と電気層110の相互関係を示す図である。電気層110は、簡略化のために、2つの従来のIPルータ111と112から構成されているものとして示されている。光層120は、ネットワークエレメントまたはノード121～125から構成されているものとして示されている。中間層130は、IPルータ112をネットワークエレメント122に結合する従来のATM/SONETシステム131を示している。層130の一部としてヘッダネットワーク132も示されているが、これは、本発明によれば、IPルータ111をネットワークエレメント121に結合している。図1には、完全な相互運用性と再構成可能性をもつ、全国規模の透過WDMをベースとするバックボーンネットワーク上のネットワーク132の配置が図示されている。

ここで強調しておきたいことは、図1の要素は本発明による実施形態の一例であり、従って、要素111は、別の実施形態では、ATMルータにすることも、スイッチにすることも可能である。

【手続補正3】

【補正対象書類名】明細書

【補正対象項目名】0023

【補正方法】変更

【補正内容】

【0023】

次に、図2を参照して説明すると、図1の光層120が詳細に示されており、この中には、本発明によれば、ネットワーク要素121～125から構成された光ネットワーク201で、高速のコネクションをセットアップするための基本的手法が含まれている。セットアップでは、そのあとに続くデータペイロード211に対して光シグナリングヘッダ210が使用される。この手法は、回線交換をベースとするWDMの利点と、パケット交換をベースとするIPテクノロジーの利点の両方を兼ね備えている。新規のシグナリングヘッダ210は、光シグナルヘッダ210の形で追加されており、マルチ波長トランスポート環境では、各波長内をインバンドで搬送される。光シグナリングヘッダ210は、パケットのソース、デスティネーション、優先度、およびパケットの長さなどの、ルーティング情報と制御情報とを収めているタグであり、データペイロード211に先行して光ネットワーク内を伝達される。各WDMネットワーク要素121～125は、光シグナリングヘッダ210を検知し、コネクションテーブル（後述する）を調べ、クロスコネクション（cross-connection）、アッド（add）、ドロップ（drop）、ドロップ・コンティニュー（drop-and-continue）などの、必要なステップを実行する。コネクションテーブルは、NC&M220とWDMネットワーク要素121～125が、チャネル221などの論理コネクションを介して継続的に通信し合うことによって絶えず更新されている。光シグナリングヘッダ210のあとに続くデータペイロード211は、コネクションによって確立された各ネットワークエレメ

ントの経路を通るようにルーティングされる（後述する）。図2の構成によれば、図2にTで示すように、光シグナリングヘッダ210とデータペイロード211の間の時間遅延を管理する必要がないのは、各ネットワークエレメント内でコネクションセットアップのために要求される短時間に対して必要な光遅延が、その間に置かれたファイバの遅延を通して各ネットワークエレメントで与えられるからである。さらに、データペイロードのフォーマットとプロトコルは、ヘッダのそれから独立している。すなわち、あるネットワークが与えられているとき、ヘッダのフォーマットとプロトコルは事前に決まっているのに対し、データペイロードのフォーマットとプロトコルはヘッダのそれと同じにすることも、異なったものにすることも可能になっている。

【手続補正4】

【補正対象書類名】明細書

【補正対象項目名】0027

【補正方法】変更

【補正内容】

【0027】

一般的に、符号化／除去モジュール132は、IPトラフィックがWDMネットワークに入ったり、出たりするインタフェースとなっている個所に置かれ、これはネットワークエレメントのクライアント・インタフェースとIPルータの間に位置している。クライアント・インタフェースは、CCIタイプにすることも、非コンプライアント・クライアント・インタフェース（non-compliant client interface NCI）タイプにすることも可能である。これらのインタフェースの個所で、ヘッダエンコーダ321は、IPシグナルがネットワーク201に送り込まれるとき、デスティネーションと他の情報を収めている光ヘッダ210を、データペイロード211の前に置く。光ヘッダ210は、インターネット311を介してIPルータ111から取得されたIPシグナルの元のIPアドレスに基づき、光変調器（後述する）によって光ドメイン内で符号化される。シグナリングヘッダリムーバ322は、クライアント・インタフェースからドロップされた光シグナルのヘッダ210を検出し、電気

的なIPパケットをIPルータ111に引き渡す。

【手続補正5】

【補正対象書類名】明細書

【補正対象項目名】0029

【補正方法】変更

【補正内容】

【0029】

図4は、第2タイプの「プラグアンドプレイ」モジュールである光エレメント410を示しているが、これは、各々のWDMネットワークエレメント121～125と関連付けられている。ここでは、説明の便宜上、エレメント121が示されている。モジュール410は、従来のネットワークエレメント回線スイッチコントローラ420と、従来のスイッチングデバイス430との間に置かれている。モジュール410は、タップファイバ経路(tapped fiber path)404～406からモジュール410へ供給された、いずれかのファイバ401～403上を伝達された各シグナリングヘッダ210からの情報を検出する。モジュール410の働きは、テーブルルックアップを高速化し、スイッチングデバイス430へのシグナリングを高速化することである。スイッチコントローラ420は、ネットワークエレメントを制御するために使用されている従来の「クラフトインタフェース(craft interface)」と同じ働きをする。しかし、この場合には、このスイッチコントローラ420の目的は、NC&M220から回線交換シグナリングを受け取り、タグスイッチコントローラ410にどのコマンドを送るべきかを、優先度に基づいて判断することである。従って、タグスイッチコントローラ410は、ネットワークエレメント回線スイッチコントローラ420から回線交換制御シグナルのほかに、シグナリングヘッダ210から引き出された情報も受信し、回線交換方式にするか、タグスイッチ制御方式にするかを統合的に選択する。スイッチングデバイス430を構成するスイッチも、スイッチングを高速化する。スイッチングデバイス430への入力経路401～403に置かれているファイバ415、416、または417に要求される遅延は、遅延がシグナリングヘッダを読み取り、テーブルルックアップ

を完了し、スイッチングを行うために要する総時間より大になるようになってい
る。概算であるが、2 kmファイバでは、処理時間は10ミリ秒になる。エレメ
ント121～125で表され、スイッチングデバイス430を含んでいるWDM
ネットワークエレメントのタイプとしては、波長アドドロップ・マルチプレク
サ (Wavelength Add-Drop Multiplexer WADM)、波長選択クロスコネク
ト (Wavelength Selective Crossconnect WSXC)、および制限波長変換能力もつ波
長インターチェンジ・クロスコネク ト (Wavelength Interchanging Crossconnect WIXC) がある。

【手続補正6】

【補正対象書類名】明細書

【補正対象項目名】0032

【補正方法】変更

【補正内容】

【0032】

各ネットワークエレメント121～125をNC&M220と結合すると、適
応性のあるルーティングプロトコルが実施される。このルーティングプロトコル
は、次のような機能を実行する。(a) ルーティングストラテジに関する通
信回線のステート、予測されたトラフィック、遅延、および容量利用率など、ネ
ットワークパラメータを測定し、(b) ルーティング計算のために測定情報を
NC&M220に転送し、(c) NC&M220でルーティングテーブルを計
算し、(d) ルーティングテーブルを各ネットワークエレメント121～12
5に配布し、各ネットワークエレメントでパケットルーティングの判断をさせる
。NC&M220は、各ネットワークエレメントからネットワークパラメータ情
報を受信し、ルーティングテーブルを定期的に更新し、そのあと(c) エレメ
ント111などのIPルータからのコネクション要求をNC&M220に転送し
、(f) NC&M220からのルーティング情報を各ネットワークエレメント
121～125に転送し、光シグナリングヘッダ210に入力させる。

【手続補正7】

【補正対象書類名】明細書

【補正対象項目名】0059

【補正方法】変更

【補正内容】

【0059】

(ルーティングプロトコル)

NG1と同規模のネットワークでは、ルーティング判断を集中化することは実現不可能であるので、このアプローチは、分散意思決定まで一般化する必要がある。階層アドレッシングとルーティングは、IPルーティングの場合と同じように使用されている。新しいコネクションが要求されると、NC&M220は、WDM経路がWDMベースのネットワーク内のそのペア（ソース、デスティネーション）に用意されているかどうかを判断する。そうであれば、パケットは、即時にその（1ホップIPレベルの）経路上に送出される。そのような経路が用意されていなければ、NC&M220は、最初のWDMネットワークエレメント用の初期出方路リンクと、新トラフィックを搬送する波長を判断する。この判断は、新コネクションの要求時に、ネットワーク内に残っているコネクションに基づいて行われる。その後、NC&M220は、該当するプロトコルを通してシグナリングを使用し、関係情報を初期WDMネットワークエレメントに転送し、シグナリングヘッダに入れられる。初期出方路リンクが判断された後、ルーティング判断の残りは、光シグナリングヘッダ情報に従って個別のネットワークエレメント（NE）で行われる。この方法によると、各スイッチングノードに置かれたルーティングテーブルと、必要とされるシグナリングヘッダ処理が、相対的に小さく保たれることが保証される。また、この方法によると、ネットワークは、スイッチングノードとネットワークユーザの点でスケーリングが容易化される。なお、複数のWDMサブネットを1つに相互接続し、各サブネットに独自のNC&Mをもたせることも可能である。

【手続補正8】

【補正対象書類名】明細書

【補正対象項目名】0062

【補正方法】変更

【補正内容】

【0062】

本発明による手法では、このルーティングプロトコルは、偏向ルーティング方式に似ている（「発明の背景」の箇所参照）。そこでは、セッションは、優先経路にたどって行くことができないとき、他の出方路リンク（優先度で見たとき）に偏向される。パケットは、継続的に偏向されることは許されていない。従来のルーティングプロトコルでは、特定数のホップの後、セッションをブロックするためにホップカウントが使用されている。新規方式では、スイッチングノードでヘッダ再生成が許されていない場合には、ホップカウント手法は使用できない。その代わりに、光シグナリングヘッダの特性（つまり、シグナリングヘッダの信号対雑音比）を調べて、パケットをドロップすべきかどうかを決定することが可能になっている。

【手続補正9】

【補正対象書類名】明細書

【補正対象項目名】0069

【補正方法】変更

【補正内容】

【0069】

図10の回路図は、図4の実施形態の詳細を示す例である。図10において、各ヘッダ検出器1010、1020、...、1050、...、または1080は、それぞれ経路1001、1002、1003、および1004上であって、デマルチプレクサ1005、1006、1007、および1008によって処理された光入力を構成している各波長からの情報を処理する。各デマルチプレクサの例は、図9の回路900に示されている。処理された情報は、各波長別にグループ化される。従って、例えば、高速メモリ1021は、ある波長が与えられているとき、ヘッダ検出器1010からのリード1011に現れたシグナルと、ヘッダ検出器1030からのリード1012に現れたシグナルと、ヘッダ検出器1050からのリード1013に現れたシグナルと、ヘッダ検出器1070から

のリード1014に現れたシグナルを、入力として受け入れる。各高速メモリ1021～1024は、内容アドレスメモリ（content-addressable memory）のように、対応するタグスイッチコントローラ1031～1034への入力となる働きをしている。各タグスイッチコントローラ1031～1034は、図4のネットワークエレメントスイッチコントローラ420から回線交換制御シグナルも受信する。各タグスイッチコントローラは、コントローラ420から与えられる回線交換コントロールと、対応する高速メモリから与えられるタグスイッチ情報のどちらかを統合的に選択し、該当制御シグナルを図4のスイッチングデバイス430に与える。

【手続補正10】

【補正対象書類名】明細書

【補正対象項目名】0079

【補正方法】変更

【補正内容】

【0079】

(d) 波長変換技術

波長変換は、経路偏向またはパケットバッファリングを不要にしてパケットコンテンションの解決に用いられる。経路偏向とパケットバッファリングのどちらにも、一連パケットのシーケンスを乱す危険がある。さらに、パケットバッファリングは持続時間だけでなく、容量にも制約があり、非透過方法が要求されることが多い。他方、波長変換は、同一経路を通るように代替波長で伝送することによってブロッキングを解決するので、同一遅延が得られることになる。図示の例では、制限波長変換能力を持つWSXCが配置されている。

【手続補正11】

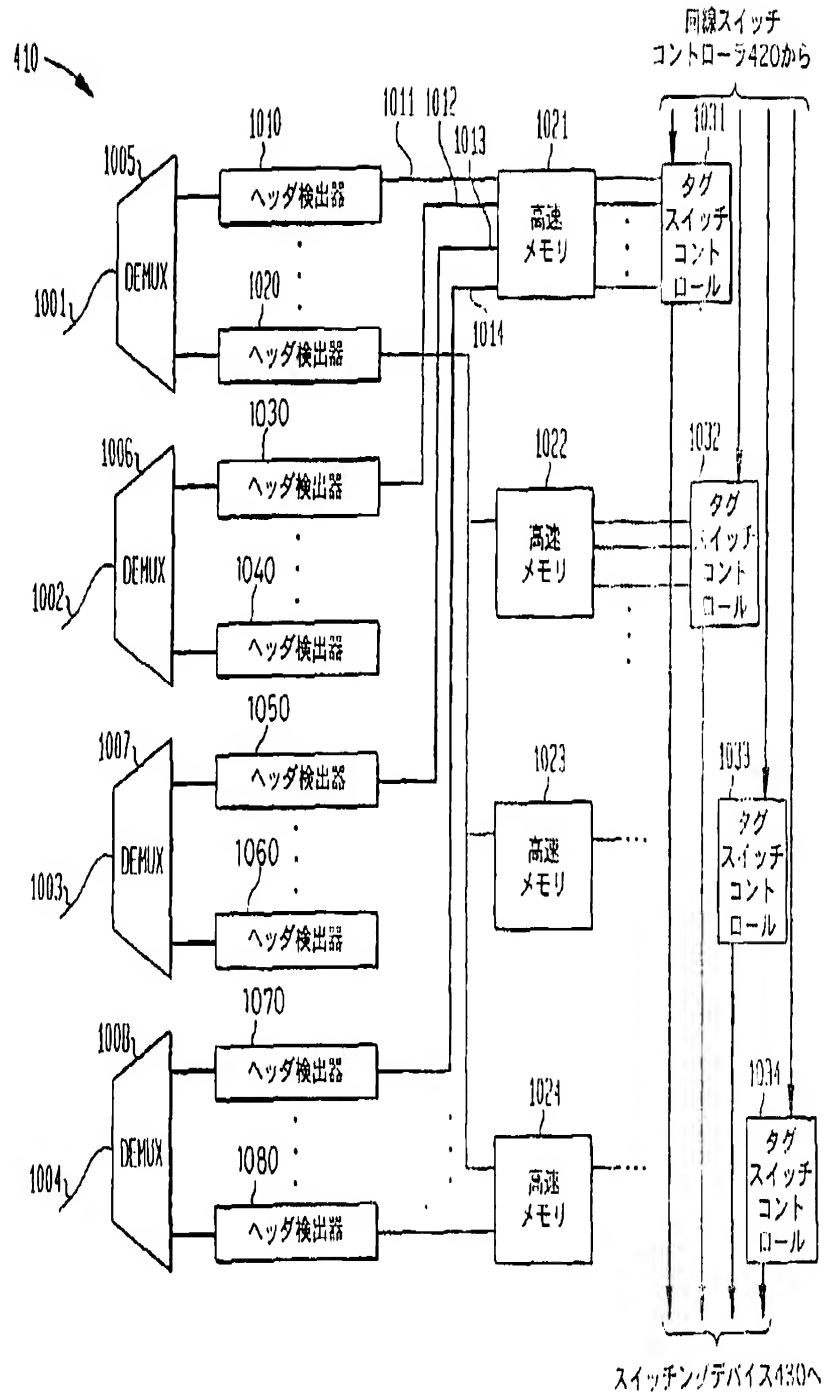
【補正対象書類名】図面

【補正対象項目名】図10

【補正方法】変更

【補正内容】

【図10】



【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/14979

A. CLASSIFICATION OF SUBJECT MATTER IPC(6) : H04J 14/02 US CL : 359/123-124,128; 370/392-393,471,474 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 359/123-124,128; 370/392-393,471,474		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,488,501 A (BARNSELY) 30 January 1996. See Fig. 1	1-2,4,7-14,16,20-22, 24
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principles or theory underlying the invention *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art *Z* document member of the same patent family	
A documents defining the general state of the art which is not considered to be of particular relevance		
B earlier document published on or after the international filing date		
L documents which may throw doubt on priority claim(s) or which are cited to establish the publication date of another citation or other special reason (as specified)		
O documents referring to an oral disclosure, use, exhibition or other means		
P documents published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search		Date of mailing of the international search report
17 SEPTEMBER 1999		19 OCT 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230		Authorized officer KINPE-MICHAEL NEGASH <i>E. P. Negash</i> Telephone No. (703) 305-4932

フロントページの続き

(72)発明者 スンジョー ヨー
アメリカ合衆国 95616 カリフォルニア
州 デービス クリーブランド ストリー
ト 635

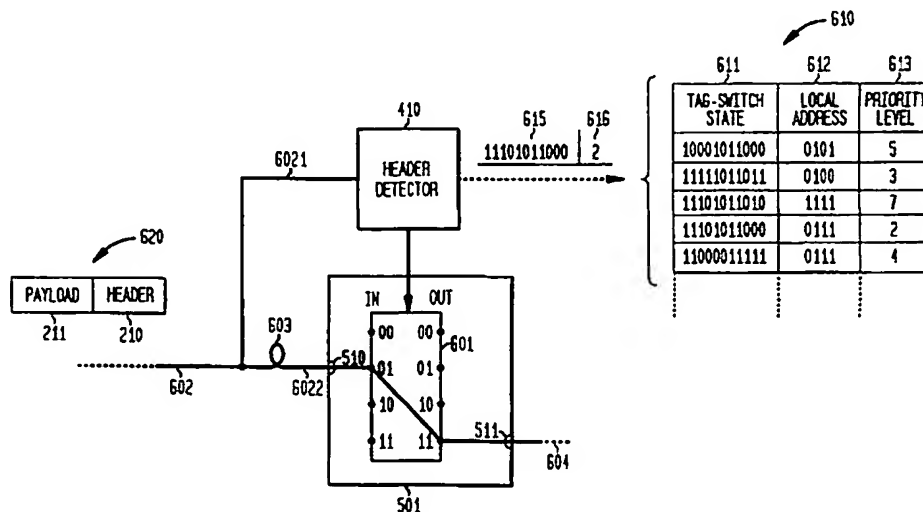
F ターム(参考) 5K002 AA06 BA02 BA06 CA14 DA02
DA03 DA05 DA09 DA13 FA01
5K028 AA11 BB03 CC02 CC05 DD04
KK03 KK32 MM05 MM12 RR01
TT02
5K030 GA03 HA08 HC01 HD03 JA01
JA11 JA14 JL03 JL07 KA05
KX20 LA07 LA17 LB05



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : H04J 14/02		A1	(11) International Publication Number: WO 00/04667
			(43) International Publication Date: 27 January 2000 (27.01.00)
(21) International Application Number: PCT/US99/14979		(81) Designated States: AU, CA, CN, ID, IN, JP, KR, MX, SG, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 1 July 1999 (01.07.99)			
(30) Priority Data: 09/118,437 17 July 1998 (17.07.98) US		Published With international search report.	
(71) Applicant: TELCORDIA TECHNOLOGIES, INC. [US/US]; 445 South Street, Morristown, NJ 07960-6438 (US).			
(72) Inventors: CHANG, Gee-Kung; 7 East Lawn Drive, Holmdel, NJ 07733 (US). YOO, Sung-Joo; 635 Cleveland Street, Davis, CA 95616 (US).			
(74) Agents: GIORDANO, Joseph et al.; International Coordinator, Rm. 1G112R, 445 South Street, Morristown, NJ 07960-6438 (US).			

(54) Title: HIGH-THROUGHPUT, LOW-LATENCY NEXT GENERATION INTERNET NETWORKS USING OPTICAL TAG SWITCHING



(57) Abstract

An optical signaling header (210) technique applicable to optical networks wherein packet (620) routing information is embedded in the same channel or wavelength as the data payload (211) so that both the header (210) and data (211) payload propagate through network elements with the same path and the associated delays. The header (210) information has sufficiently different characteristics from the data payload (211) so that the signaling header can be detected without being affected by the data payload, and that the signaling header can also be removed without affecting the data payload. The signal routing technique can overlaid onto the conventional network elements in a modular manner using two types of applique modules. The first type effects header encoding and decoding at the entry and exit points of the data payload into and out of the network; the second type effects header detection at each of the network elements.